AD A061153

LEVEL (12)

ARI TECHNICAL REPORT
TR-78-A24

Development of a Model Tank Gunnery Test

by

George R. Wheaton and Paul W. Fingerman
AMERICAN INSTITUTES FOR RESEARCH

and

G. Gary Boycan
ARI Unit Training and Evaluation Systems Technical Area

August 1978

Contract DAHC-19-76-C-0003

Prepared for

U.S. ARMY RESEARCH INSTITUTE
for the BEHAVIORAL and SOCIAL SCIENCES
5001 Eisenhower Avenue
Alexandria, Virginia 22333

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

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

JOSEPH ZEIDNER
Technical Director (Designate)

WILLIAM L. HAUSER
Colonel, US Army
Commander

Research accomplished
under contract to the Department of the Army

American Institutes for Research

NOTICES

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

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

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR-78-A24	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DEVELOPMENT OF A MODEL TANK GUNNERY TEST		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER AIR-55800-8/77-FR
7. AUTHOR(s) George R. Wheaton, Paul W. Fingerman (AIR) and G. Gary Boycan (ARI)		8. CONTRACT OR GRANT NUMBER(s) DAHC 19-76-C-0003
9. PERFORMING ORGANIZATION NAME AND ADDRESS American Institutes for Research 1055 Thomas Jefferson Street, NW Washington, DC 20007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q763743A773
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Armor School, Directorate of Training Developments Fort Knox, KY		12. REPORT DATE August 1978
		13. NUMBER OF PAGES 132
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) US Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Research monitored technically by G. Gary Boycan of the Unit Training & Evaluation Systems Technical Area of the Army Research Institute.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Performance testing
Armor crew testing
Model tank gunnery test development

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Research was undertaken to develop a model live-fire test that can be used to evaluate tank crew proficiency in neutralizing targets. The model test takes into consideration different types of target engagements, the behaviors of the individual crew members that are required and the practical constraints associated with the use of main gun ammunition for testing purposes.

Existing descriptions of M60A1AOS gunnery objectives were reviewed and updated to reflect current US Army Armor School doctrine. The revised set of

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) objectives provided a definition of the possible ways that the tank system could neutralize targets with the 105mm gun, the coaxial machinegun and the .50 caliber machinegun. The objectives also made explicit the behaviors required on the part of individual crew members during each engagement. Given a definition of the performance domain, two steps were undertaken to aid in the selection of objectives or exercises that would constitute the model gunnery test. First, all objectives were systematically examined to identify possible groups or families of objectives that were similar to one another in terms of the underlying crew behaviors involved. Families were further analyzed to generate estimates of the extent to which performance on one gunnery objective was predictive of performance on other objectives in the same family. These results, when coupled with other considerations such as the number of main gun rounds available for testing, were then used to identify a representative sample of test engagements that would yield information about crew performance on the entire domain of gunnery objectives.

Issues involved in the generation and interpretation of performance test scores were also addressed. Four test purposes were identified: crew qualification, prediction of combat effectiveness, skill diagnosis and crew motivation. For each purpose alternative scoring strategies were examined and appropriate procedures defined.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of the many individuals who assisted in the research reported herein. Major R. M. Bosserman and Mr. Robert K. Bauer of the Directorate of Training Developments, U. S. Army Armor School at Fort Knox aided greatly in the conceptualization of Armor training and testing. Staff Sergeant Rodney Caesar, of the Main Battle Tank Branch, Weapons Department, Armor School, deserves our special thanks for his painstaking and insightful review and revision of the gunnery domain data matrix. His special efforts in confirming the detailed behavioral descriptions of each objective, and in determining conformance with emerging doctrine were critical.

A number of contractor and subcontractor personnel also made valuable contributions during the effort. Dr. John A. Boldovici and Mr. William C. Osborn of HumRRO and Dr. Andrew M. Rose of AIR assisted in conceptualization of the project, played major roles in the development of item selection criteria and prepared several interim project documents from which material was adapted for this report. Other personnel who assisted in conducting studies and analyzing and interpreting data included: James H. Harris, Charlotte L. Heinecke, Ronald E. Kraemer, Richard E. O'Brien, Roy C. Campbell and Jack R. Reeves. Many of these individuals were also involved in the IDOC-1 project which supplied the baseline for the tank gunnery domain developed herein.

3/4 Blank

DEVELOPMENT OF A MODEL TANK GUNNERY TEST

BRIEF

Requirement:

To develop a model tank gunnery test that can be used to measure crew proficiency in neutralizing targets. The model test takes into consideration different types of target engagements, the behaviors of the individual crew members that are required and the practical constraints associated with the use of main gun ammunition for testing purposes.

Procedure:

Existing descriptions of M60A1AOS gunnery objectives were reviewed and updated to reflect current U.S. Army Armor School doctrine. The revised set of objectives provided a definition of the possible ways that the tank system could neutralize targets with the 105mm gun, the coaxial machinegun and the .50 caliber machinegun. The objectives also made explicit the behaviors required on the part of individual crew members during each engagement.

Given a definition of the performance domain, two steps were undertaken to aid in the selection of objectives or exercises that would constitute the model gunnery test. First, all objectives were systematically examined to identify possible groups or families of objectives that were similar to one another in terms of the underlying crew behaviors involved. Families were further analyzed to generate estimates of the extent to which performance on one gunnery objective was predictive of performance on other objectives in the same family. These results, when coupled with other considerations such as the number of main gun rounds available for testing, were then used to identify a representative sample of test engagements that would yield information about crew performance on the entire domain of gunnery objectives.

Issues involved in the generation and interpretation of performance test scores were also addressed. Four test purposes were identified: crew qualification, prediction of combat effectiveness, skill diagnosis and crew motivation. For each purpose alternative scoring strategies were examined and appropriate procedures defined.

Findings:

Revisions in the original pool of 225 gunnery objectives resulted in a net addition of 41 objectives, increasing the total to 266. One hundred and twelve different crew behaviors were identified in specifying the performance requirements associated with the various objectives.

Computer analyses of the behaviors that objectives had in common with one another permitted the objectives to be grouped into 16 families. Within each of these, objectives were then ordered based on the degree to which they were representative of the family. Sampling of families on a proportional basis resulted in the identification of 28 exercises for inclusion in the model gunnery test.

This sample of test exercises satisfied a number of criteria considered critical to the design of an effective test. First, the test contained at least one highly representative objective from each major family, thereby providing a basis for inferences about the quality of performance in each family and by extension the entire gunnery domain. Second, the exercises covered the range of tactical and environmental conditions under which engagements may occur. Finally, the test exercises required the crew to perform most of the 112 crew behaviors associated with gunnery. Only 10 behaviors were not included in the model test, and of these, nine occur rarely in the domain of 266 objectives.

The performance data that are generated by the test exercises can be used to satisfy multiple test purposes. For crew qualification, a criterion-referenced scoring approach was adopted. Crews must perform a specified percentage of main gun and machinegun engagements to the standards required. The actual percentages used in labeling crews qualified, marginally qualified or unqualified are set so as to minimize potential classification error. A similar approach applied to selected engagements is also detailed for estimating crew effectiveness in combat. For motivational purposes, normative scoring approaches are suggested, including use of existing point-scoring systems. Finally, for training diagnosis, the use of performance profiles is recommended. Such profiles can be used to isolate specific deficiencies in crew performance as well as to identify shortcomings in the training program.

Utilization of Findings:

Crew proficiency in the use of tank weapons is a major goal of gunnery training. The model test provides training managers with the instrument needed to manage such training more effectively. The analytic procedures and the sampling strategy which evolved may also be applied in the design or revision of similar tests for other tank systems in the inventory. Finally, the model test provides a systematically defined set of criterion measures that can be used to evaluate promising techniques for simulation-based testing of tank crews.

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

As specified in FM 17-12 (1977), gunnery training and testing for both active and reserve component tank crews consist of three logically sequenced stages of development. Initially, the focus is on the individual and the acquisition of pre-gunnery skills. The focus then shifts to the training of crews in the use of tank weapons to neutralize a variety of targets. Finally, attention centers on tactical training in which tank sections/platoons acquire skill in distributing fire during platoon battleruns.

During the second stage of the program, in which crew training is emphasized, the training content is organized within eight firing tables, each of which consists of several training or testing exercises. Early tables provide the gunner and tank commander with the rudiments of sound technique. Later tables involve the whole crew, training them to function as a team in using the tank's weapon systems, in compliance with latest doctrine, to neutralize targets under a variety of engagement conditions. As the culmination of this stage of training the U.S. Army annually requires crews to demonstrate specified levels of weapons proficiency while firing Table VIII for record.

The exercises comprising Table VIII are constantly undergoing revision in response to changing weaponry and doctrine. In each instance, however, the exercises have been selected and developed on the basis of competent opinions and the judgment of experienced armor personnel who, realizing that exhaustive testing of crews is impossible because of resource constraints, have attempted to distill the essence of gunnery into a manageable set of test exercises. Similarly, they have devised a variety of scoring systems to provide for differentiation among crews.

Recently, however, there has been a growing interest in revising both the exercises comprising Table VIII and the strategy for its administration and scoring. At issue are a number of test-construction questions concerning test content, the cost-benefit of alternative test strategies, the amount of data required for decision making and the interpretability of test scores. In the final analysis the resolution of these and related issues hinges on two considerations. What aspects of crew performance should Table VIII address? What uses are to be made of the resulting test data? Both issues are explored below in the course of developing a rationale for Table VIII testing. The discussion parallels that presented in more detail elsewhere (Wheaton, 1977).

TEST CONTENT OF TABLE VIII

As described above, tank gunnery training is sequenced into a series of stages. Progression through the program is dictated, at least implicitly, by a number of hurdles placed at the end of each training phase. In order to proceed to crew gunnery training, for example, the individual crewman must first pass a series of go/no-go test exercises comprising the Tank Crew Gunnery Skills Test (TCGST). Similarly, upon completion of the crew gunnery phase of training, each crew must demonstrate a specified level of proficiency on Table VIII before participating in platoon training exercises. What aspects of gunnery performance should Table VIII address? The answer depends on how one defines the domain of performance subsumed under the rubric of tank gunnery.

At its most basic level, gunnery must include "marksmanship"--the ability of the crew to neutralize targets under a variety of engagement conditions within a reasonable amount of time. The content and method of scoring most current Table VIII's indicate that they certainly strive to measure marksmanship. In fact, the crew gunnery standards specified in FM 17-12 (1977) reflect this basic kind of performance for M60 series tanks as follows:

- Given an M60-series tank moving 12-15 mph, main gun loaded and laid no more than 15^{o} off target, range and ammunition indexed (1,000 meters/HEAT or 1,600 meters/APDS), engage an armor-type target, using battlesight, within 5 seconds during daylight, and within 10 seconds during darkness under artificial illumination.* As a minimum, a target hit should

* When white light illumination is used, daylight scoring times will be used.

be obtained within 10 seconds during daylight and within 15 seconds at night at ranges to battlesight range.

- o Given an M60-series tank moving 12-15 mph, main gun loaded and laid no more than 15° off target, engage an armor-type target, using precision fire, within 10 seconds during daylight, and within 15 seconds during darkness under artificial illumination.* As a minimum, a target hit should be obtained within 15 seconds during daylight and within 20 seconds at night from battlesight range to 2,500 meters.

- o Given a moving or a stationary hulldown, M60-series tank, the crew will adjust, fire a second round, and obtain a target hit within 5 seconds of a first-round miss.

- o Given an M60-series tank, the crew will engage a troop-type target, at a range not exceeding 1,600 meters, using the caliber .50 machinegun, or 900 meters using the coax machinegun, within 5 seconds during daylight, and within 10 seconds during darkness under artificial illumination.*

- o Given an M60-series tank and a range card previously made by the crew, the crew will re-position the tank at night, and, using range card data, hit targets within 5 minutes of reaching the referenced position.

- o Given an M60-series tank moving no more than 15 mph and a fire command, the crew will engage a lightly-armored vehicle or aircraft with the caliber .50 machinegun within 10 seconds during darkness under artificial illumination.*

With respect to basic marksmanship, therefore, proficiency testing in Table VIII involves the systematic application of the standards cited above to hit/miss and engagement time raw data. A main gun engagement in which a hit is not obtained is regarded as a failure and scored as zero (0). When a main gun hit is obtained, the engagement is regarded as a success and scored as one (1) if: the first round is fired within

---

* When white light illumination is used, daylight scoring times will be used.

(15)

the prescribed time, and the hit is secured within the designated time, and the second round, if needed, is fired within the interval allowed. Machinegun engagements would be similarly scored in terms of target hits and engagement times.

Other current Table VIII's attempt to go beyond basic marksmanship (USAREUR Reg. 350-704, 1976), providing for the inclusion and scoring of such tasks as the selection of the most threatening target from among multiple targets, choice of an appropriate method of engagement, and target acquisition, as well as more peripheral components such as effective use of existing terrain for cover and concealment. In these Table VIII's interest lies in testing such factors as tactical decision making and planning in addition to marksmanship. Three problems are associated with this approach.

The first problem lies in the inextricable intermingling of the marksmanship and decision-making components in the scoring of crew performance. If, for example, a crew fails a given exercise, is it because they can't shoot or because of faulty decision making, such as engaging the targets in the wrong order? When a single engagement is used to measure performance on a compound criterion, failure cannot clearly be attributed to one or the other of the components of the criterion. In keeping with the sequential nature of the training program, two separate tests of crew proficiency would seem desirable. The first would focus on marksmanship, and provide test results unambiguously determined by marksmanship ability. Once adequate marksmanship is demonstrated, the second test would deal with the tactical decision-making component. Even if the second test involves marksmanship, failure may be unambiguously related to decision making since satisfactory marksmanship ability has been previously demonstrated.

This is not to say that Table VIII testing of marksmanship should be accomplished in the sterile and unrealistic environment of the shooting gallery. Similarly, the argument is not intended to preclude measures of crew tactical decision making when such can be obtained. Rather, it is being suggested that unconfounded estimates are needed of a crew's ability to "put steel on target." For the purposes of this research it is assumed that Table VIII must provide such estimates by focusing on an evaluation of crew marksmanship.* Tactical decision-making components can also be incorporated and scored (e.g., by including multiple-target engagements), but their inclusion must not interfere with the primary objective of assessing the crew's marksmanship proficiency.

The second problem with the compound emphases of current Table VIII's is related: By allowing a crew to select, for example, the method of engagement, the usefulness of obtained marksmanship data on these exercises may be devalued. If a crew decides to engage a target using an inappropriate (i.e., contrary to doctrine) or less than optimal method, little information about marksmanship is gained whether the crew hits or misses. If the target is missed, was it because the crew lacked proficiency in the method of engagement selected (i.e., failed due to a lack of marksmanship proficiency) or because the method selected was inadequate for the engagement situation? Even if a hit was achieved, no information about the crew's ability to carry out a specific engagement method of interest would be obtained, unless they fortuitously selected that method. If the choice of engagement method were restricted, information regarding hits would be more useful, but this would provide a very poor test of decision making. This leads us to the third problem.

This problem is equally compelling. To assess tactical decision making adequately one would probably argue for a test of crew proficiency

*In fact, the test of crew marksmanship described in this report could be administered immediately upon completion of Table VI, the last gunnery training table which strictly addresses marksmanship. In this case, the use of comparable test items in any later table would be unnecessarily redundant. Future organization of sets of gunnery tables might, therefore, be designed to test marksmanship (e.g., a new kind of Table VII, based on this project), and then to train (e.g., a new kind of Table VIII) and test (e.g., a new Table IX) tactical decision making as distinct from marksmanship.

which occurs in the most realistic tactical setting imaginable within existing technological and cost constraints. Such a table would probably consist of engagements in which crews were free to conduct the exercises as they saw fit. That is, who fired, the method of engagement used, the range to target accepted, the fire control instrument selected, the order in which targets are engaged, etc., would be left to the discretion of each participating crew. Such freedom would be desirable and entirely consistent with a test of tactical decision making.

Were such exercises to comprise a Table VIII test of crew gunnery, however, one would be hard pressed to control the test, thereby insuring that each crew fired the same exercise in the same manner. The delivery of a prescribed and unvarying set of exercises is crucial if one is to characterize crew marksmanship across a broad range of conditions. Without such rigid specification of the conditions of engagement, crews would presumably elect to engage targets using only those techniques with which they were most proficient, whether the situation called for those methods or not. Not only would this bias contribute to variation in the test item, but it would also make difficult the testing of backup or other methods of engagement lying within the capabilities of the weapon system. These three problems can be minimized in a Table VIII which deemphasizes tactical decision making and concentrates instead on the basic aspects of gunnery associated with marksmanship.

PURPOSES OF TABLE VIII

Table VIII must be designed in both content and scoring procedures to serve multiple purposes. The four that are discussed below include: crew qualification, skill diagnosis, prediction of combat effectiveness, and crew motivation. Much of the ensuing discussion is paraphrased from a related report (Wheaton, 1977).

18

Crew Qualification. One reason given for annual exposure of tank crews to Table VIII is crew qualification. In one sense qualification implies eligibility to participate in subsequent training exercises at the platoon or company level. In another it suggests that a qualified crew is one which is combat ready, either capable of entering into combat or, once in combat, capable of succeeding. In still another sense to qualify means to be fit, to exhibit a required degree of ability. This last definition is basic and, in fact, underlies the two preceding usages.

To qualify in tank gunnery a crew must reach or exceed a certain level of proficiency. Within the marksmanship context of tank gunnery, the concept of "a certain level of proficiency" assumes two vitally important but different meanings. First, on any given engagement, crew performance must equal or exceed generally agreed upon standards of proficiency. Second, the crew must demonstrate its capacity to meet such standards on a large proportion if not on all of the engagements constituting the realm of marksmanship. A similar dualism is noted in the TCGST in which the performance of each task is required to equal or exceed specified standards, and this level of performance is demanded for every task in the test.

Qualification on the Table VIII described in FM 17-12 is based on the crew's performance in neutralizing 11 standard targets which are engaged during the day (VIIIA) and again at night (VIIIB), but in a different sequence. Twelve of the 22 targets involve main gun engagements, each of which is worth a maximum of 100 points; the remaining 10 machine-gun engagements are weighted half as heavily, each being worth a maximum of 50 points. Additional points for the use of terrain and conservation of ammunition raise the total possible score to 2000. In order to qualify, a crew must accumulate a minimum of 1400 out of the possible 2000 points (i.e., it must score 70% or better).

But with respect to the concept raised earlier of "a certain level of proficiency", what does a score of 1400 or 70% mean? Perhaps the answer can best be given by considering three hypothetical crews who qualify with the fairly similar scores of 1400, 1484, and 1468 as shown in Table 1. Obviously, the three crews are anything but similar. The first crew (1400 points) qualifies in spite of an apparent inability to neutralize targets at night with the main gun. The second crew (1484 points) qualifies without getting a single point for a machinegun engagement. The third crew (1468 points) also qualifies but fails to perform a single main gun engagement within the specified crew standard time limits! It is clear that the point scoring system which is used permits qualification even though individual engagements are not performed to standards, and even though ability to perform large segments of tasks in the marksmanship domain is in doubt.

Qualification on the USAREUR Table VIII (USAREUR Reg. 350-704, 1976) is determined on the basis of crew performance during 16 engagements. Ten of these are fired during the day (VIIIA) and six are fired at night (VIIIB). Eleven involve the main gun and five are fired with the machineguns. In order to qualify, a crew must satisfy four separate criteria:

1. at least seven of the 11 main gun targets must be hit (64%), and
2. at least five of the 11 main gun targets must be hit with the first round, and
3. opening time must average 20 seconds or less over the 16 targets, and
4. all machine gun point targets must be hit and all area targets must receive 3/5s coverage.

This approach, which represents an instance of criterion-referenced testing, provides a summary score having more meaning than the point score discussed above. We know that qualified crews are able to hit all point machinegun targets, provide at least 3/5s coverage on machinegun area targets, hit at least 64% of their main gun targets (45% with the first round) and open fire within an average of 20 seconds. This kind of information is much more directly interpretable in terms of "a certain level of proficiency."

20

Table 1

Hypothetical Performance of Three Qualified Crews
on USAARMS Table VIII A & B
(Adapted from Wheaton, 1977)

	Engagement	Maximum Score	Crew 1 Score		Crew 2 Score		Crew 3 Score	
1.	Main Gun, Battlesight	100	1st rnd. hit	100	1st rnd. hit	100	2nd rnd. hit & slower than crew standard	64
2.	" "	100	"	100	"	100	"	64
3.	" "	100	"	100	"	100	"	64
4.	" " (TC)	100	"	100	"	100	"	64
5.	" Precision	100	"	100	"	100	"	64
6.	" "	100	"	100	"	100	"	64
7.	.50 Cal. Point	50	Hit	50	Miss or too much time	0	Hit	50
8.	Coax Suppressive	50	"	50	"	0	"	50
9.	.50 Cal. Area	50	"	50	"	0	"	50
10.	Coax Point	50	"	50	"	0	"	50
11.	Coax Area	50	"	50	"	0	"	50
12.	Main Gun, Battlesight	100	2nd rnd. hit	58	1st rnd. hit	100	2nd rnd. hit & slower than crew standard	64
13.	" "	100	Miss or too much time	0	"	100	"	64

Table 1 (Cont'd.)

	Engagement	Maximum Score	Crew 1 Score		Crew 2 Score		Crew 3 Score	
14.	Main Gun, Battlesight	100	Miss or too much time	0	1st rnd. hit	100	2nd rnd. hit & slower than crew standard	64
15.	" " (TC)	100	"	0	"	100	"	64
16.	" Precision	100	"	0	"	100	"	64
17.	" "	100	"	0	"	100	"	64
18.	.50 Cal. Point	50	Hit	50	Miss or too much time	0	Hit	50
19.	Coax Suppressive	50	"	50	"	0	"	50
20.	.50 Cal. Area	50	"	50	"	0	"	50
21.	Coax Point	50	"	50	"	0	"	50
22.	Coax Area	50	"	50	"	0	"	50
	Use Terrain	200		200		200		200
	Save Ammo	100	6 rnds.	42	12 rnds.	84	0 rnds.	0
	Total	2000		1400		1484		1468

22

From the foregoing it is clear that a criterion-referenced testing approach is needed in order to determine whether crews are qualified in tank gunnery. Raw speed and accuracy measures would be evaluated against specified performance standards on an engagement by engagement basis. The crew's resultant qualification score would be interpreted as having some sort of absolute meaning such as degree of mastery of all of the engagements constituting marksmanship. This score would be interpreted totally independent of the performance of any other crew.

Skill Diagnosis. Another purpose in determining crew weapons proficiency in Table VIII is to identify those exercises on which crew performance is not up to standards. When focusing on individual crews such information may suggest specific weaknesses and lead to the prescription of particularly germane remedial training. When these data are aggregated across units they may permit diagnosis of deficiencies (and strengths) in the training system itself.

The test used to support crew qualification can also be used for diagnostic purposes. Raw score speed and accuracy measures would be examined in terms of the standards specified for each engagement. A variety of crew and crewman procedural variables could also be utilized. Performance would be characterized in terms of a profile of the individual engagements with specific weaknesses being keyed to particular combinations of the conditions defining the engagements.

Prediction of Combat Effectiveness. Another reason, often implicitly given for assessing crew weapon proficiency on Table VIII, is prediction of crew combat effectiveness. Although usually couched in such terms as combat readiness or preparedness, the focus seems to be on probability of success in combat. The usual approach is to identify engagements that are highly probable in combat or in some sense are especially important or critical. The persistent and pervasive notion is that if one can determine how well crews perform such exercises within the context of Table VIII, one can then make predictions about their performance in combat.

23

At best, exercises included within Table VIII may permit limited inferences of this type. For example, if crews can perform to standard on particularly combat-critical engagements, they may be more likely to succeed than crews that perform fewer of the key exercises satisfactorily. Actual success in combat, however, will also be dependent on other types of proficiency such as tactical decision making and the ability of crews to function as members of the platoon.

For this application exercises might be singled out from, or added to, those comprising the qualification test described above. This subset of "combat-related" engagements could then be scored separately.

Crew Motivation. A final purpose for evaluating crew weapon proficiency is to use the obtained scores to differentially reward crews on the basis of their performance. Such motivational devices generate healthy competition and appear to have a positive effect on morale.

The scoring approach which is relevant in this case is based on norm-referenced measurement. A given crew's performance is scored and evaluated in terms of how it compares relative to the performance of other crews. In a later section of this report, consideration is given to the kinds of performance which should be used as a basis for comparison, how that performance should be scored, and what aggregation procedure is to be used to generate a summary score for each crew.

In summary, it is essential that tank crews receive training and practice that maintains their gunnery proficiency at high levels. As part of this program it is also essential that crews be examined periodically to determine their level of competence, and to diagnose aspects of their performance in need of further enhancement. In addition, there is interest in attempting to forecast their combat readiness, and in differentially rewarding truly superior gunnery performance.

To accomplish one or more of these ends a variety of Table VIII's has been developed. Different commands have used alternative approaches in developing the tables, tailoring the component engagements to their own probable combat situations and using a variety of scoring procedures. Each of these approaches has its strengths and each has its weaknesses.

Consideration of the USAARMS and USAREUR Table VIII's serves to highlight the alternative tests which have been developed for the evaluation of gunnery proficiency. One uses engagements having characteristics which are readily specifiable and controllable in a testing sense. The other employs engagements which are more flexible but which are harder to manage from a testing point of view. One uses a scoring system which stresses fine discrimination among crews on a relative basis. The other tends to characterize performance in terms of what crews can and cannot do. In the final analysis the relative utility of either Table VIII (or of any other table for that matter) depends upon the uses that are to be made of obtained test data. The discussion above has explored four such uses, each of which has implications for selection of engagements, specification of performance measures, and development of scoring procedures.

PURPOSE OF THE RESEARCH

The research described in this report represents the first phase of a larger program which is concerned with the development of cost-effective techniques for evaluating tank crew weapons proficiency. As described above, the first phase is concerned with the recommendation of candidate tasks for inclusion in a new or modified Table VIII gunnery test, together with associated scoring procedures. Given a valid live-fire criterion test known to yield reliable measures of gunnery ability, a subsequent phase of effort will examine the feasibility of using simulation techniques as cost-effective alternatives to live-fire testing.

In undertaking the design of a model Table VIII a great deal of attention was given to the issues of test content and purpose. As a consequence, the objective finally adopted was development of a Table VIII which would serve multiple purposes. However, in striving toward this goal it was realized from the outset that multiple purposes of the type described above could probably never be perfectly served by a single testing instrument. In the final analysis it was decided that this goal might be most nearly realized by designing the Table VIII to be optimal for crew qualification. Given the basic set of exercises needed to support this kind of test, the other uses could be accommodated by incorporating different but interrelated scoring approaches, and, in some cases, adding a few exercises sensitive to issues other than qualification.

In the remainder of this report the test development approach is described and the resulting model Table VIII is presented in detail. In the next section the concept of a gunnery domain of performance is introduced. This notion is used to explore the need for, as well as alternative approaches to, the sampling of gunnery performance. Subsequent major sections then deal with the actual specification of test items for inclusion in Table VIII, an elaboration of alternative scoring systems and a discussion of considerations during implementation of the model Table VIII. In each of these presentations an effort has been made to avoid technical exotica; when *unavoidable*, they are treated in appropriate appendixes.

ITEM SAMPLING STRATEGIES

The primary purpose of the model Table VIII is to determine whether a tank crew can satisfactorily perform those gunnery tasks which might reasonably be expected of them. The first step in refining what aspects of gunnery would or would not be covered in such a test lay in the decision to emphasize crew marksmanship as opposed to tactical decision making. The second refinement was to define the nature and range of the activities subsumed under the term "marksmanship." For purposes of the present effort the area or domain of performance termed marksmanship was to include all ways in which crews could neutralize targets within the constraints of current Army doctrine and within the capabilities of the M60A1AOS weapon system.

Each of the possible ways in which crews can neutralize targets may be referred to as a job objective, that is, a component of the more general job of "tank gunnery." Developing the pool of job objectives permits one to operationally define the content domain of the test. Having specified the domain of concern, it is then fair to ask whether a crew can successfully accomplish each of the constituent job objectives.

The goal in all test development efforts is to answer this question by selecting a set of items from the domain that best predicts performance on the entire domain within existing cost and other constraints. (If cost were not important, there would be no need to sample items. All subjects could be tested on the entire domain of job objectives.) The task of selecting objectives from a domain for use as test items can be accomplished in many ways. Ultimately though, the decisions about how to choose items will be made on the basis of considering cost and the purpose of the test. The methods for test development used in this project were chosen after considering the following possible bases for selecting items from the domain of gunnery job objectives:

1. Random sampling,
2. Frequency of performing the objectives on the job,
3. Performance data,

4. Generalizability of the objective, or "communality" with other objectives in the domain, and

5. Criticality of the objectives.

The following discussion of these considerations has been liberally adapted from material presented in an earlier working paper (Boldovici, Wheaton, & Boycan, 1976).

RANDOM SAMPLING

Since the goal of gunnery training is to promote mastery of all objectives in the job domain, a mastery test could be devised by randomly sampling objectives from the domain. The reasoning for this approach follows the justification for random sampling of subjects in preparing experimental designs. Just as the performance of a randomly constituted sample of subjects in an experiment will approximate the performance of the (untested) entire population, the performance of a crew on a random sample of job objectives will be an estimate of their performance on the entire domain. The obvious advantage of such an approach is in test security. Very large numbers of alternate forms of the test could easily be devised, even to the point of having a different form of the test for each group of test-takers. The use of random sampling also would eliminate problems associated with "teaching the test."

The main disadvantages of random sampling (or of other sampling methods that would result in different performance tests for different crews) are two-fold. The first is that because of range facility limitations it would not be possible to administer all the test exercises that might be sampled. The second is the difficulty which this approach creates when addressing the other testing purposes. Because each crew would receive a different set of items, skill diagnosis, while feasible on a crew-by-crew basis, could not readily be extended to an evaluation of the overall training program. To the extent that predictions of combat effectiveness

27

depended on testing certain key or benchmark objectives, one would be at the mercy of the draw. Finally, relative appraisal of crews would be difficult if not impossible because of the lack of a common set of test items.

One could of course approach random sampling from a slightly different direction in order to deal with some of these problems. The alternative would be the random drawing of a single set of exercises. The primary advantage of this approach is its simplicity--drawing the single random sample is straightforward, and is easily performed anew at any time in the future when the domain might be reconstituted. Further, unlike the random alternative forms approach described above, comparison of crews is straightforward.

In spite of the apparent attractiveness of one or more variants of random sampling, its chief disadvantage is that it lacks power. It ignores information potentially available about the relationships among objectives, and therefore its ability to predict performance for the entire domain is weakened. This corresponds to the distinction in the development of experimental designs between pure random sampling, and blocked sampling. If the population to be sampled can be subdivided on the basis of a priori information or variables which may be relevant to the desired performance data, the statistical power of the sampling design may be improved by dividing the sample appropriately prior to sampling. As will be seen below, a great deal of information can be assembled about the domain which, when used to assist in the selection of job objectives for testing, can greatly improve the efficiency of the model Table VIII.

FREQUENCY OF PERFORMANCE

Frequency of task performance is an easily obtained and sometimes interesting measure. The problem is that equally tenable cases can be made for including high-frequency and low-frequency tasks in tests. How the measure would be used in test design is, therefore, not clear.

The case for including high-frequency tasks in tests hinges on the relation between frequency of task performance and importance: If task Y is the most frequently performed part of Job X, then task Y is in a literal sense "important" to the performance of Job X. If one accepts the propositions that frequency of performance is indicative of task importance, and that item selection should be made on the basis of task importance, then one would design tests that measured frequently performed tasks.

The case for including low-frequency tasks in tests is just as easily made. The most frequently performed tasks may be the least difficult of job tasks, may not be generalizable with respect to other tasks in the domain, and may not even be necessary for effective job performance. And inasmuch as a high frequency of task performance on the job may guarantee eventual mastery of the task, one could argue that testing resources should not be expended for evaluating such tasks; that is, assuming that testing resources should be expended in ways that will yield maximum information about the extent to which a job has been mastered, then the expenditure of resources on measuring the performance of tasks whose mastery is highly probable is unwarranted. One would therefore design tests that measured infrequently performed tasks.

After exploring both of the positions presented above, we concluded that frequency of task performance has no clear implications for design of a Table VIII primarily concerned with crew qualification.

Were the primary thrust one of forecasting crew combat effectiveness, the first argument presented above (i.e., the more frequently performed tasks are the more important tasks) would be of relevance. However, it would be difficult to generate estimates of the relative frequency of performance of specific gunnery objectives since the context would be in the future, in combat, and specific to a particular theater of operations. Nevertheless, such information might be of use in pinpointing gunnery objectives that could serve as benchmarks of combat gunnery, thereby supporting estimates of crew effectiveness.

PERFORMANCE DATA

Traditional test-development methods could be used for developing a domain-referenced test. Task difficulty and performance variability

30

would play a role in item selection, producing a test capable of making fine discriminations among the scores of various crews. Such an approach would begin by collecting repeated measures of actual live-fire performance on all items in the gunnery domain. The data would then be subjected to an item analysis, and sets of items would be identified for inclusion in Table VIII which best predicted total scores on the entire domain. The particular set selected would depend on the trade-off in predictive power and the cost of administering each set.

There are several problems with this approach. The first is that available resources simply do not permit obtaining the necessary repeated measures on the multitude of gunnery objectives presumably comprising the domain. Second, where this approach would support skill diagnosis and the relative assessment of crews, it would provide estimates of combat effectiveness only to the extent to which key or benchmark objectives were found in the predictor set of items. Finally, while this traditional approach is ideally suited to performance measurement or estimation in a norm-referenced sense, the primary concern of the model Table VIII is crew qualification, an issue to be resolved with a criterion-referenced test. In this criterion-referenced approach, one addresses the binary issue of mastery for each objective. Arriving at an estimate of crew qualification requires concentrating on comprehensiveness, or when comprehensiveness cannot be achieved (due to unacceptable costs), on item "generalizability."

GENERALIZABILITY

There is a problem posed in the preceding section. On the one hand, item generalizability is precisely what empirical methods such as item analysis are designed to establish. On the other hand, the use of item analysis with empirically collected performance data must be ruled out as too expensive. Nevertheless, generalizability remains an essential criterion for item selection if performance on the test is to be predictive of performance on the entire domain of job objectives. The solution to this problem is to develop indirect methods for estimating the generalizability among job objectives, and to employ these estimates in place of performance-based measures.

CRITICALITY

An additional way of prioritizing job objectives for inclusion in a test of tank gunnery is on the basis of criticality. For example, one approach might be to describe the criticality of objectives in terms of the threat which the target situation represents for the crew. The selection of test items on this basis would be defensible were one primarily interested in using Table VIII to predict effectiveness or success in combat. Crews would be tested on those objectives representing the most threatening conditions, that is, those on which their own survival hinged. However, the use of a criticality sampling criterion could well be at odds with other test purposes, including crew qualification and skill diagnosis.

Even if one were to opt for a Table VIII optimized for the prediction of combat effectiveness, one would still face a formidable challenge in translating the concept of criticality into actual weights that could be used to prioritize objectives. Expert ratings of criticality are notoriously unreliable (Smode, Gruber, and Ely, 1962). Comparison-type ratings, for instance of target threat, usually are more reliable, but are very costly to generate when the number of items to be compared is large.

An alternative approach would simply be to redefine the domain of marksmanship objectives, including only those objectives, for example, that reflected primary methods of engagement, or main gun engagement of tank targets. A significant penalty would, of course, be exacted for doing so. The exclusion of non-critical objectives would seriously hamper efforts to determine a crew's mastery of the larger, overall domain of gunnery.

SUMMARY

Three of the sampling strategies discussed above could be used to prioritize objectives for inclusion in a Table VIII: frequency, criticality, and generalizability. (The fourth approach, based on empirical and traditional test-construction methodology, is not practical in the present situation.)

Assuming that all three weighting schemes were equally plausible, which they are not, how would they be used to prioritize objectives? Consider for instance, the following example in which an eight-objective domain is portrayed and criticality, frequency, and generalizability weights have been established for each objective--higher values representing more desirable items.

Job Objective	Generalizability	Criticality	Frequency	Sum
1	8	3	1	12
2	7	4	2	13
3	1	1	8	10
4	2	2	7	11
5	3	8	3	14
6	4	7	4	15
7	6	5	5	16
8	5	6	6	17

Suppose one wanted to construct a two-item gunnery test from the candidate objectives. Were one to draw items to satisfy the generalizability criterion, objectives 1 and 2 would be chosen. Similarly, in maximizing criticality and frequency one would choose objectives 5 and 6, and 3 and 4, respectively. Each of the three tests developed in this manner would satisfy one selection criterion optimally (e.g., 5 and 6 would be the most critical objectives) but would fail to optimize the others (e.g., objectives 5 and 6 are neither the most generalizable nor the most frequent objectives). One could of course, elect to use a composite weight, such as the sum of the three values generated for each objective. In this case he risks failing to represent any single criterion optimally (e.g., he chooses objectives 7 and 8).

It is clear, therefore, that testing purposes must drive the selection of appropriate sampling strategies, and that when considered simultaneously, these may lead to the selection of sets of test items which at best partially overlap and at worst are mutually exclusive. The result will either be a series of different but optimal tests, or a single test which is in no sense optimal.

As already stated, the primary purpose of the model Table VIII under development in this project was crew qualification. In light of the problems discussed above, it was decided to use generalizability as the primary way of prioritizing objectives for selection of test items. Other testing purposes were to be served by modifications of scoring strategies or the addition of supplementary objectives. In the following sections the method of carrying out this item sampling strategy is described.

SELECTING OBJECTIVES FOR A MODEL TABLE VIII

Given rationales for the purposes of Table VIII testing, and identification of the general performance domain of primary interest, attention turned to the problem of selecting those items which would constitute the Table VIII test. As described below, four steps were involved in the item-selection process. First, a specific domain of performance was identified which served to define the crew job-objectives comprising tank gunnery. Second, the behavioral details of crew members' roles in each objective were elaborated. Third, relationships among job objectives comprising the domain were examined as a basis for selecting objectives for Table VIII. The fourth and final step was to draw the sample of test items.

DEFINITION OF THE DOMAIN OF TANK GUNNERY

Initial specification of the domain of tank gunnery was based on a list of 225 job objectives reported by Kraemer, Boldovici, and Boycan (1975). In keeping with the emphasis on crew marksmanship, these objectives defined all possible ways that a variety of targets could be neutralized with the main gun, coaxial machinegun, and .50 caliber machinegun weapons of the M60A1AOS tank system.

The description of each objective in the pool of 225 consisted of three components: a conditions statement, a task statement, and a standards statement. A typical objective was:

> Given (a) a stationary M60A1AOS tank with the main gun battlesighted with SABOT or HEAT, (b) an operational gunner's day periscope, and (c) a stationary tank or light-armored vehicle target that is visible at less than 1100 meters without artificial light at day or night; the gunner will open fire within 7 seconds of the alert element of the tank commander's (TC's) command, and neutralize the target within 12 seconds, using no more than two rounds.

Each of the objectives was written in the form of the example cited above. The task and standards components of the objectives were, in all cases,

". . .open fire within ____ seconds of the alert. . ., and neutralize the target within ____ seconds, using no more than ____ rounds."

Objectives for inclusion in the pool were formed by combining levels of 11 conditions associated with hypothetical engagements and attaching these combinations to the task and standards statements. The 11 conditions were: crew member firing, weapon, fire-delivery method, firing vehicle motion, target motion, target type, target visibility, target range, day vs. night firing, fire-control instrument, and ammunition. The levels within specific conditions are listed in Table 2 .

Several sets of revisions were made in the 225 job objectives defined by Kraemer, Boldovici, and Boycan (1975). The revisions were undertaken to reflect emerging gunnery doctrine, and in response to a series of reviews by the project staff and by the staffs of the Weapons Department of the U.S. Armor School and the Directorate of Training Developments, Fort Knox, Kentucky. When deciding whether to modify, delete, or add an objective the reviewers were guided by the capabilities of the M60A1AOS weapon system and current doctrine and not by what kinds of engagements form the basis of training or by what engagements are typically fired. For example, the original pool of objectives included precision engagements fired from moving tanks. Current doctrine dictates, however, that the firing vehicle will come to a brief halt prior to engaging targets in the precision mode. Thus, in order to update the domain in accordance with this doctrinal change, precision engagements fired on-the-move were redefined to be fired after moving to a halt. Other changes in doctrine led, for example, to the deletion of coaxial machinegun engagements fired in the precision mode, as well as main gun engagements fired by the tank commander (TC) at night with the aid of his metascope. Still other objectives were added to the original pool to reflect possible encounters with aircraft targets, moving targets in range-card-lay-to-direct-fire engagements, etc. Collectively, these revisions resulted in a net addition of 41 job objectives, increasing the pool from 225 to 266. The complete set of job objectives is included in Appendix A.

36

Table 2. ENGAGEMENT CONDITIONS AND LEVELS OF CONDITIONS USED TO DEFINE THE JOB-OBJECTIVE POOL

(Adapted from Kraemer, et al. 1975)

CONDITIONS

Crew Member	Weapon	Firing Mode	Firing Vehicle Motion	Target Motion	Target Type	Target Visibility	Day or Night	Fire Control Instrument	Ammunition	Target Range
1. Gunner	Main Gun	Battlesight	Stationary	Stationary	Thin-skinned Vehicle (TSV)	Visible	Day or Night	Range-finder, Day	Sabot or Heat	≤1600 meters
2. Tank Commander	Coax	Precision	Moving	Moving	Tank (TNK) or Light Armored Vehicle (LAV)	Visible using artificial illumination	Night	Range-finder, Infrared (Metascope)	Hep	≤1000 meters
3.	.50 Caliber	Non-precision	Moving-to-a halt		TSV or Crew Served Weapon	Not visible		Tank Commander's Periscope, Day Light	Beehive	500-4400 meters
4.		Range Card			Bunker or Crew Served Weapon			Tank Commander's Periscope, Infrared	Coax 7.62 mm	500-3200 meters
5.		Range Card-lay-to-direct-fire			Troops			Gunner's Periscope, Day Light	.50 Caliber Machinegun	≤4400 meters
6.					LAV or Crew Served Weapon			Gunner's Periscope, Infrared		500-1600 meters
7.					LAV			Telescope		≤3200 meters
8.					LAV or Crew Served Weapon or Troops			Infinity Sight		500-1200 meters
9.					Aircraft			Auxiliary Fire Controls		≤900 meters
10.					All Targets					≤2300 meters

BEHAVIORAL DESCRIPTION OF THE DOMAIN

Once the job objectives forming the gunnery domain had been defined in terms of system capabilities and doctrinal constraints, the next step was to characterize each job objective in behavioral terms. This step was undertaken to provide a basis for estimating the extent to which performance on any given objective would be generalizable to or predictive of performance on any other job objective or on the domain as a whole. Ultimately, this generalizability could serve as a basis for item selection in the model Table VIII.

In traditional test development exercises the problem of item generalizability is handled by actually obtaining repeated measures of performance on each candidate test item from a large sample of subjects. Item analyses are then performed using the empirically obtained performance data, and subsets of items that best predict performance on the entire domain, for a given cost, are identified. Because available resources did not permit collection of repeated measures of live-fire performance on each of the 266 candidate job objectives, a rational approach was devised for establishing the generalizability from one objective to another. The validity of this approach rests on the assumption that the more "task elements" or "behavioral steps" an objective has in common with other objectives, the greater will be the communality among the objectives, and the greater will be the probability that performance on the one objective is predictive of performance on the others. For example, if it can be shown that firing battlesight at 1000 meters has more behavioral steps in common with (1) firing battlesight at 1200 meters than with (2) firing the .50 caliber machinegun at 1000 meters, it is assumed that the battlesight-1000-meter engagement will be more predictive of engagement (1) than of engagement (2). Furthermore, given that communality between a task and the entire domain could be derived, it is assumed that a task with high "domain communality" would be more representative of (or predictive of) performance on the entire domain than one with low communality.

To provide the data needed for the analysis of behavioral commun-

38

ality among objectives, detailed descriptions of the activities of each crew member (i.e., tank commander, gunner, driver, and loader) were developed for each of the 266 tank gunnery job objectives. Members of the project staff who were subject-matter experts listed the behavioral elements for each crew member involved in each job objective. Examples of behavioral elements are: "Loader announces 'Up'", "Gunner indexes HEP", and "Gunner levels bubble."

Certain "system state" assumptions had to be made early in the analysis in order to achieve consistency in specifying which behavioral elements were and were not included in each objective. If the weapons were assumed to be loaded, for example, then the behavioral elements involved in a target engagement would be different from those that would be involved if the weapons were assumed to be unloaded. Specifying the ground rules under which the analysis would occur began with two sets of assumptions:

1. If the target to be engaged is visible without artificial illumination, then the firing vehicle will be in the battlesight mode, in which all weapons are loaded and SAFED, and the firing switches are OFF. The turret power, computer, primary direct fire control instruments, main gun stabilization system are ON; and the caliber fifty machinegun rate of fire selector switch is in LOW. Depending on type of target being engaged, the appropriate ammunition is indexed into the computer and the corresponding range is indexed into the rangefinder (e.g., 1100 meters for HEAT and 1600 meters for SABOT).

2. If the target to be engaged is not visible under any circumstances or is visible only if artificially illuminated, then the firing vehicle will be in the range card mode, in which a range card has been made out for HEP ammunition, the main gun is not loaded, and the brakes are locked. The machineguns are fully loaded and all weapons are SAFED. The turret power, computer, primary fire control instruments, main gun stabilization system, and the firing switches are OFF.

34

Assumptions had to be made in addition to those cited above. These other assumptions, which pertain to weapons, ranging, fire control instruments, firing vehicle and target motion and range, are presented in Appendix B.

The system-state assumptions discussed above pertain to conditions that exist during and shortly before target engagement. Assumptions also had to be made about the termination of target engagements.

Recall that the job objectives do not specify whether a first- or second-round hit is to be achieved. The possibility of accomplishing a job objective in more than one way (i.e., neutralizing a target with the first or second round) presented a problem for the present analysis, because the number and kinds of behavioral elements in any job objective will differ depending on whether or not a first-round hit is achieved. The obvious way to solve this problem is to assume either that all engagements will terminate in a first-round hit, or that all engagements will result in a first-round miss followed by fire adjustment and refiring. Because neither of these assumptions is consistent with the results of real target engagements, we were reluctant to adopt one or the other; and since no a priori compromise was apparent, two data matrices were generated--one which identified the task elements involved in accomplishing each objective via a first-round hit, and another which identified the task elements involved in accomplishing each objective by firing, missing, adjusting fire, firing again, and achieving a second-round hit.

To facilitate the behavioral descriptions, a job objective - behavioral element data matrix was generated as shown in Figure 1. Arabic numerals were assigned to each of the 266 job objectives and were entered along the left margin of a large matrix. The analysts then listed the task elements for each crew member involved in the first job objective and coded these by entering numbers across the top of the matrix. A "1" was then entered under each task-element number in the row for Objective #1, indicating that performance of the objective required performance of each of the behavioral elements in the row. The behavioral elements for Objective #2 were then specified. Any new elements -- ones that were not

40

Job Objective Number	Behavioral Element Number												
	1	2	3	4	5	6	7	8	9	10	11	12	114
1	1	1	1	1	1	0	0	0	0	0	0	0	
2	1	1	1	1	1	1	1	1	0	0	0	0	
3	1	1	1	1	0	0	0	0	1	1	1	1	
4													
5													
6													
7													
8	1	1	1	1	1	0	0	0	1	1	0	1	
266													

Figure 1. Job objective-behavioral element data matrix.

included in Objective #1 -- were coded and their numbers added to the top row of the matrix. The same was done for Objective #3 and for the remaining 263 objectives. An example of the specification of behavioral elements is given for Objective #8 in Table 3. When coded and entered into the eighth (objective) row in the data matrix (Figure 1), these behavioral elements would describe a pattern of "1"'s and "0"'s uniquely associated with Objective #8. A "1" indicated that a behavioral element occurred during the crew's performance of that objective, while "0" entries indicated that given elements were not involved.

The number of behavioral elements required to describe the 266 objectives using the first-round hit assumption was 114. An additional 21 elements were associated with subsequent-round engagements in which fire was adjusted by means of either burst-on-target technique (eight additional elements), a standard range adjustment (six additional elements), or a subsequent fire command (seven additional elements). As indicated earlier, which of these three methods (if any) should be appended to the description of the basic objectives would ultimately depend upon post hoc information about a specific engagement. That is, did the first round miss and if so, what method of adjustment was appropriate? In light of this a priori indeterminacy, and the fact that inclusion of a subsequent method of engagement would at most increase the pool of behavioral elements by only 7% (8/114), it was decided to analyze the relationship among objectives based on the primary pool of 114 elements associated with firing of the first round. The list of behavioral elements used to describe the objectives is attached as Appendix C*.

---

*As with the set of job objectives, the list of task elements was revised several times with the assistance of the Weapons Department of the Armor School at Fort Knox, and this process continues as doctrine and practice are revised. Two of the 114 elements were recently deleted, after most of the analyses reported below had been performed; it has been determined that these deletions would have had only very minor impacts on the analyses, which therefore were allowed to stand. The two deleted elements are indicated in Appendix C.

Table 3. SAMPLE JOB OBJECTIVE DESCRIBED IN TERMS OF TASK ELEMENTS AND ELEVEN GENERAL CONDITIONS

Job Objective No. 8 — Crew Member: Gunner — Weapon: Main Gun — Firing Mode: Battlesight

Vehicle Motion: Moving — Target Motion: Moving — Target Type: Tank or LAV

Target Visibility: Visible — Target Range: <1600 Meters — Day/Night: Day or Night

Fire Control Instrument: Telescope — Ammunition: Sabot or Heat

TANK COMMANDER BEHAVIORS:

1	TC Announces "Gunner"
16	TC Announces "Battlesight"
45	TC Announces "Moving"
44	TC Announces Target Description
25	TC Lays Gun for Direction
90	TC Announces "Fire" or "At My Command, ... Fire"

GUNNER BEHAVIORS:

36	Gunner Turns on Main Gun Switch
42	Gunner Selects Sabot or Heat Reticle
47	Gunner Announces "Identified"
74	Gunner Applies Lead in Direction of Target Apparent Motion
77	Gunner Lays Rangeline Leadline at Center of Base of Target
94	Gunner Times Shot
98	Gunner Announces "On the Way"
100	Gunner Fires Main Gun

LOADER BEHAVIORS

14	Loader Unlocks Ammo Ready Rack
26	Loader Selects Sabot or Heat
31	Loader Places Main Gun Safety in Fire Position
33	Loader Announces "Up"

DRIVER BEHAVIORS

8	Driver Maintains Steady Rate of Speed
9	Driver Maneuvers Tank for Firing
10	Driver Announces Adverse Terrain Conditions

ANALYSIS OF RELATIONSHIPS AMONG OBJECTIVES

The basic issue to be addressed in analyzing the 266 job-objective by 114 behavioral-element matrix was whether the entire domain consisted of highly similar objectives, or of several subdomains, each containing highly similar objectives but differing from other subdomains. A unitary domain, defined in terms of relatively high communality among the behavioral elements comprising all the job objectives, would lead to the assumption of generalizability among the entire pool of objectives. Sampling would, therefore, proceed from among the entire set of 266, the selected items generalizing to the domain as a whole. If, however, subdomains or "families" of objectives were identified in which objectives were homogeneous within each subset and heterogeneous among subsets, a different sampling strategy would be used. In this case, since objectives would be assumed to generalize best to others in the same subdomain, a stratified sampling plan would be in order. Test items would be selected from each homogeneous subset. Representation of the domain would be ensured by including test items in the model Table VIII from each subdomain of objectives.

Procedure. In order to determine whether the domain of gunnery objectives was unitary or consisted of a number of subdomains, a technique known as cluster analysis was employed. This technique, when applied to the kind of data contained in the job-objective/behavioral-element matrix, is used to search for a series of partitions which define groups with similar patterns of data. In this particular instance groups of job objectives were sought which had similar patterns of behavioral elements. The analysis was conducted by subjecting the 1, 0 data contained in the objective-element matrix to the BMDP3M cluster analysis program contained in the Biomedical Computer Programs package (Dixon, 1975). A more detailed discussion of the analytical procedure and of the interpretative process is presented in Appendix D.

Results. The cluster analysis indicated that the domain of tank gunnery as defined in this study might best be characterized in terms of sixteen relatively homogeneous clusters or families of job objectives. The clusters, which correspond fairly well to a rational grouping of objectives based on major conditions of engagement, are summarized in Table 4, together with the number of job objectives contained in each.

The first cluster contains 30 job objectives in which the gunner uses the main gun in the battlesight method of engagement to engage targets with SABOT or HEAT ammunition. The analysis indicated that objectives in this cluster could be broken down more finely as a function of firing vehicle motion (stationary, moving to a halt, and moving), but this finer breakdown only slightly increased the homogeneity of the resulting three subclusters over that of the larger cluster which, therefore, was retained on practical grounds. The 30 job objectives in the cluster have the same value on 94 out of 114 behavioral elements.

The second cluster is similar to the first since it includes main gun SABOT/HEAT battlesight engagements; however, for these 12 objectives, the tank commander fires the weapon. This cluster can also be broken down by firing-vehicle motion, but the resulting gain in homogeneity is small. This cluster was formed by adding one anomalous task to the statistically derived cluster, since the outlying task fit better here than elsewhere. The relatively high uniformity (98 out of 114 elements) suggested that addition of the outlier objective to the 11-objective cluster defined in the computer solution was reasonable.

The third cluster in Table 4 includes the 16 objectives describing gunner main gun precision engagements using SABOT or HEAT. This cluster could be further broken down on the basis of target motion (stationary or moving), as well as on the basis of firing-vehicle motion. Again, however, the resulting gain in homogeneity is small. The 16 objectives in the composite cluster have the same values on 97 out of 114 elements.

45

Table 4. CLUSTERS OF JOB OBJECTIVES

Cluster*	Crew Member	Weapon	Method	Fire Control Instrument	Ammunition	Cluster Size
1	G	MG	BS		S/H	n = 30
2	TC	MG	BS		S/H	n = 12
3	G	MG	P		S/H	n = 16
4	TC	MG	P		S/H	n = 8
5	G	MG	P		H/B	n = 24
6	TC	MG	P		H/B	n = 12
7	G	MG	RCL		H/B	n = 10
8	G	MG	RCL		S/H	n = 6
9	TC	MG	RCL		S/H	n = 2
10	TC	MG	RCL		H/B	n = 3
11	G	CX	NP		CX	n = 56
12	G	CX	RCL		CX	n = 13
13	TC	CX	NP		CX	n = 24
14	TC	CX	RCL		CX	n = 6
15	TC	.50	NP	TPI	.50	n = 15
16	TC	.50	NP	TPD	.50	n = 29
						266

* Tank motion, target motion, target type, visibility, day or night, fire control instrument, and range are usually mixed within clusters.

Abbreviations:

- G = Gunner
- TC = Tank Commander
- MG = Main Gun
- CX = Coax Machinegun
- S/H = Sabot/Heat
- H/B = Hep/Beehive
- .50 = .50 Caliber Machinegun
- BS = Battlesight
- P = Precision
- RCL = Range Card Lay to Direct Fire
- NP = Nonprecision

The fourth cluster, consisting of eight tank commander objectives, is similar to the third. It also formed from two subclusters. Again, the overall cluster has relatively high uniformity (101 out of 114 elements).

The fifth cluster is similar in content to the third, with the exception of a change in ammunition to HEP/BEEHIVE. The 24 objectives have similar values on 86 of the 114 elements.

The sixth cluster involved the 12 objectives defining tank commander precision engagements using HEP or BEEHIVE. Objectives in the cluster have 88 elements in common.

The seventh and ninth clusters described in Table 4 were unusual and will be discussed at the end of this section.

The eighth cluster consisted of the six objectives in which the gunner fired SABOT or HEAT using *the range-card-lay-to-direct-fire or the* range-card methods of engagement. It was formed from a five-objective cluster, plus the addition of one anomalous isolate (c.f. the second cluster). The six objectives in the cluster have similar values on 104 elements.

The tenth and final main gun cluster includes tank commander range-card engagements using HEP and BEEHIVE. Objectives in this cluster have 102 elements in common.

The analysis segmented machinegun engagements into six clusters, defined primarily by the crew member firing and the method of engagement. Cluster 11 includes the 56 job objectives in which the gunner engages targets using the coaxial machinegun in the non-precision mode. The cluster analyses indicated that this large cluster might be more finely characterized if broken down by target motion and target type (area vs. point targets), but with only a small gain in homogeneity. The clustered objectives have 90 elements in common.

(47)

The twelfth cluster is made up of the 13 coax range-card-lay-to-direct-fire exercises fired by the gunner. It formed from two smaller clusters and one isolated range card objective. The objectives comprising the overall cluster had similar values on 98 elements out of 114.

The thirteenth and fourteenth clusters are tank commander clusters corresponding to clusters 11 and 12. They share 96 and 106 elements respectively.

The final two clusters are tank commander .50 caliber clusters which are differentiated on the basis of the sight (fire control instrument) used to fire the weapon. Cluster 15 includes the 15 objectives which employ the tank commander's infrared sight. Its objectives have 99 out of 114 elements in common. Cluster 16 includes the objectives which employ the tank commander's daylight sight. Its objectives share 98 elements.

The two clusters for which discussion was postponed above actually were never defined as such by the computerized cluster analysis. Cluster 7 (gunner firing HEP or BEEHIVE using range-card-lay-to-direct-fire) and cluster 9 (tank commander firing SABOT or HEAT using the same method of engagement) were formed entirely ad hoc, from objectives which were almost randomly dispersed throughout the remaining main gun job objectives. These objectives were moved out of the clusters into which the cluster analysis originally placed them because their contents (both in terms of conditions and behavioral elements) were so different; indeed, in most cases they were remote outliers as far as the computerized solution was concerned. Thus, in some sense these two clusters comprise "other" categories. Fortuitously, however, these "other" categories turned out to be relatively homogeneous with respect to content, and to fill in the gaps in the basic solution matrix, defined by crew member, weapon system, and mode of fire. Further, objectives within these constructed clusters were found to have relatively large numbers of elements in common (90 and 110 respectively) indicating that they were just about as homogeneous as the remainder of the clusters. The conclusion

48

is that for these 12 exercises, most of the flaws in the algorithm (discussed in Appendix D) coincided to prevent cluster formation. However, the fact that the solution was directly referrable to the underlying behavioral elements permitted a straightforward adjustment.

In summary, results of the cluster analysis indicate that the set of 266 job objectives do not simply characterize a single job of "tank gunnery", but rather approximately 16 jobs as described above. These differ primarily in terms of the weapon involved, the method of engagement, and the crew member who fires. Further distinctions involve ammunition and fire control instruments. For purposes of testing to determine if a crew can perform the "tank gunnery" job, therefore, it will be necessary to test whether they can perform each of the 16 distinct components of that job. The domain of job objectives in Appendix A has been organized according to the 16 clusters described above.

SAMPLING OF JOB OBJECTIVES

The sixteen sets of job objectives emerging from the cluster analyses suggested that a stratified sampling approach be used when selecting test items for the model Table VIII. In implementing this strategy two issues needed resolution: The first concerned the appropriate number of items to sample; the second was to select the best way of sampling using as the primary criterion the construct of generalizability within sets of objectives.

Sampling Strategy. The most straightforward approach to determining the number of items for the model Table VIII would be to select 16 objectives (or multiples of 16), one (or more) for each of the salient clusters in the domain. However, there are obvious tradeoffs among the number of test items, testing efficiency, and ultimately, test cost. As a rough guideline to the resource/cost constraints current versions of both the USAARMS and USAREUR Table VIII's are somewhat longer than 16 exercises. Thus it was felt that resources currently available could support more than 16 exercises, especially since adding exercises would only improve the accuracy of qualification decisions. The second reason for considering more than 16 items was to insure approximately equal testing power for each distinct set of objectives. Were only 16 exercises used, then each family in the domain would be represented by only one test item; however, some families (e.g. gunner, coax, non-precision with 56 objectives) are much larger than others (e.g., tank commander, range card, SABOT/HEAT with only two objectives). If one were to try to estimate performance on one family of 56 objectives and on another of two with a single test item for each, there would necessarily be less predictive power for the larger family. It seemed appropriate, therefore, to attempt to have approximately equal testing power for each of the families within the bounds of a practical test.*

Therefore, the stratified sampling approach was implemented by adopting a sampling rate proportional to the size of each family. Based on this rule, a family with twice as many objectives as another family would be represented by twice as many test exercises. In order to have a Table VIII of reasonable size, a sampling fraction of .1 was used. The number of exercises in each family was divided by 10 and rounded to the nearest integer number; this then was the number of exercises to be selected from that family. For example, the largest family (56 job objectives) would be represented by six exercises, the second largest (30) by 3, and so on. On this basis two clusters containing two and three objectives respectively (tank commander range-card-lay-to-direct-fire using SABOT/HEAT and using HEP/BEEHIVE), were too small to be represented at all. Since the two were similar in content and differed only in terms of the ammunition used, they were collapsed into a single set for sampling, and one item was selected.

Armed with guidelines as to the number of objectives to sample from each cluster, the next step was to develop an index of generalizability to use in actually drawing samples of objectives from each family.

---

* To the extent that the model Table VIII exceeds future resource availability it may be reduced to the minimum of 16 objectives using rules described below. This represents an alternative cost-effectiveness tradeoff.

As argued previously, generalizability in this context is based on the assumption that the more behavioral elements in one objective that are contained in the other (and vice versa), the more predictability there will be from either objective to the other. Thus, any objective which includes all of the elements involved in any other objective in the cluster would be ideal. Unfortunately, this is never the case, since most changes in engagement conditions from objective to objective produce corresponding changes in the element structure, and therefore no objective in any family has all of the elements which ever occur in each of the remaining objectives. An approximation to this ideal would be to select the objective which contains a majority of the behavioral elements that are ever involved in the family. By this criterion there are usually several objectives, particularly in the larger families, that would qualify. But this criterion amounts to simply counting elements--the objective in any family which has the most elements involved would necessarily have the most elements in common with any other objective in the family. Therefore, in developing a generalizability index it was decided to weight these primitive frequencies by how often the element actually occurred in the family. This was done to emphasize the relative importance of elements within the family. Similarly, to the extent that an element is frequent in a particular cluster and infrequent in the domain, it is even more important that such an element be represented in items selected for testing, since such elements describe behaviors which are unique to the particular kinds of engagements represented in the cluster. An additional weighting dealt with this issue. The index resulting from this compound weighting procedure was termed the "generalizability index" ($\Sigma F_i^2/D_i$) and was used to order objectives within families to establish sampling priorities. For further details of the generalizability index see Appendix E.

Sampling Procedure. Using the derived index $\Sigma F_i^2/D_i$, selection of the most generalizable objectives from each cluster was straightforward. Within each of the 15 strata to be sampled from (recall that two tank commander clusters were collapsed into one stratum) the index was computed for constituent objectives, and these were ranked from high to low in terms of generalizability. Based on an examination of rank values, the job objective with the highest index value was selected from each cluster. In clusters from which more than one objective was to be chosen, the additional items were drawn according to the same rule. The objectives ranking second, and third, and so on, were selected until the quota for a given cluster was reached.

The prototype set of objectives resulting from application of this procedure revealed several features needing additional consideration. First, two objectives involved either the use of BEEHIVE ammunition or a moving aircraft target. Doctrinal and resource constraints uncovered during the course of the project suggested that it would be difficult to fire these engagements under normal Table VIII testing conditions. Nevertheless, summary rejection of such objectives seemed unwarranted; it was decided to retain them in a model Table VIII (provided they met other sampling criteria discussed below) since quality of test content rather than ease of implementation was the primary concern.

Further inspection of the prototype test exercises raised questions about the adequacy with which various engagement conditions were represented. In striving for maximum generalizability based on the ranked index values, various engagement conditions were over - or under-represented. For example, in five of the seven strata from which two or more objectives were to be drawn, tied ranks were encountered. In each instance the tie occurred between day and night versions of what otherwise were virtually identical objectives. Inclusion of both exercises, as demanded by the initial sampling procedure, served to restrict the variety of engagement conditions represented. As another example, in spite of the inclusion of day/night twins, a preponderance of nighttime engagements resulted. Ten of the 13 main gun engagements were at night as were 12 of the 15 machinegun exercises. A similar imbalance resulted with respect to the fire-control instrument used to conduct the engagement. Nine of the 13 main gun exercises involved the use of secondary or backup sights instead of primary ones, and in only one instance was the gunner's day periscope represented.*

In order to deal with these imbalances the original sampling strategy was modified in a number of ways. The basic thrust of the revised sampling procedure was to: 1) obtain a balance between day and night engagements for both the gunner and the tank commander; 2) use day and night engagements when multiple objectives are drawn from a given cluster; 3) provide for the representation of more primary and fewer backup methods of engagement; and 4) maximize generalizability to the extent possible within these constraints.

The four tank commander main gun clusters were addressed first. One objective was needed to represent each of these clusters, and it was drawn by identifying that exercise which ranked first in terms of the generalizability index. In each case the highest ranking objective consisted of a *nighttime engagement*; consequently balancing was needed. One of the clusters, comprised of range-card-lay-to-direct-fire engagements, required use of a nighttime objective since it contained no day engagements. Therefore, the day/night balancing was undertaken using the other three clusters. The day twins of the highest ranking objectives ranked second, second, and third in generalizability. A balance was

*These imbalances were not unexpected, and did not decrease the basic validity of $\Sigma F_i^2/D_i$. This index describes the generalizability of objectives in terms of the communality of behavioral elements and does not address conditions of engagement. Thus, there is no problem with the index, but rather there are concerns with face validity and user acceptance. For example, while doctrine permits many types of secondary engagements, some are very undesirable relative to their primary counterparts. Thus, the emphasis on behavioral element generalizability must be tempered by these other considerations.

struck by selecting the two second-ranking day twins and using them in conjunction with the night engagements representing the other two clusters.

Attention turned next to the five gunner, main gun clusters. The procedure was similar but somewhat more complicated in this case because of the need to balance across both day and night and primary and secondary sights, and to insure when more than one objective was to be sampled from a cluster that both day and night exercises were chosen. In all, nine objectives were to be drawn to represent the five clusters. First, the highest ranking objective in each cluster was identified. Two of these came from range-card-lay-to-direct-fire clusters and were immediately adopted as night engagements. The most generalizable objective in each of the three remaining clusters was represented by a pair of day/night twins tied with the highest rank. One objective from each pair was eventually to be chosen. In addition to picking these three, four other engagements were to be chosen to round out sampling quotas for the main gun clusters. Selection of these items was more involved than the choice of the two range-card exercises.

The first step was to decide how the overall sample of gunner main gun exercises should be apportioned between day and night engagements. A nearly equal five/four split seemed desirable, but there was no compelling reason to choose five night and four day exercises instead of the reverse. Accordingly, samples were eventually drawn that represented both of these alternatives. The second step was to settle upon a balance between backup and primary methods of engagement. Since generalizability dictated three objectives that involved backup methods, the decision was made to include only primary methods when selecting the remaining objectives.

The third step was to develop rules for those instances in which two or more objectives were to be drawn from the same cluster. The notion of simply proceeding with the second-and third-most generalizable objectives was rejected for reasons cited earlier--the approach leads to an over-representation of night and backup engagements. As an alternative it was decided to examine the fabric of each cluster in question to determine whether a finer characterization of the cluster's content was possible. This essentially meant working backward through the cluster tree diagrams previously described.

The clusters in question fortuitously broke down into subclusters or segments whose number equalled the number of objectives to be drawn. To take advantage of this circumstance the objective having the highest cluster generalizability was drawn first and the segment to which it belonged was noted. That segment of the cluster was then excluded from further consideration. Within-segment generalizability index values were then used to draw the subsequent objectives when these were needed.* The sampling rule was to draw the highest ranking primary-method-of-engagement objective from each segment as well as the highest ranking daytime engagement fired with a preferred sight.

The sampling procedure resulted in the identification of six objectives from among which to draw three required gunner battlesight items; four precision, SABOT/HEAT engagements from which to draw two items; and four precision, HEP/BEEHIVE objectives from which to choose two items. Combinations of items were selected from these pools that reflected the various balancing criteria.

The final stage of sampling involved the selection of machinegun engagements. The procedure was essentially the same as used in sampling main gun items. As much of a balance as possible was desired between day and night engagements for the tank commander and gunner; and when possible both types were to be drawn from any given cluster or segment. When multiple items were required the first was drawn on the basis of cluster generalizability, while the remainder were drawn in terms of

*The three objectives drawn on the basis of cluster generalizability also had the highest segment generalizability within the segment to which they belonged.

engagements having the highest segment generalizability.* Nine objectives were identified with which to represent the two gunner clusters. Twelve were found which could be used to represent the four tank commander clusters. Because of the absence of day/night twins, the criterion to balance day and night conditions sharply curtailed the number of combinations which were possible.

Sampling Results. Before the model Table VIII could be compiled from the machinegun, tank commander/main gun, and gunner/main gun components, the latter had to be given further scrutiny. A total of 16 objectives was drawn, nine of which were required for representation of the gunner's performance with the main gun. Twenty-four alternative combinations of nine objectives were identified which satisfied the various sampling criteria. The pool of items and the resulting item combinations are presented in Appendix F.

In an attempt to select that one combination which best represented this portion of the domain, a number of subsidiary analyses were undertaken. First, a listing was made of the behavioral elements underlying each candidate set of objectives. The lists did not differ. In each case the 24 alternative combinations provided for coverage of the same 67 behavioral elements. Since alternatives couldn't be differentiated in terms of behavioral elements, coverage of engagement conditions was examined next. In 16 of the 24 cases the gunner's day periscope was represented more frequently than his infrared periscope. In the remaining eight cases (i.e., see objectives 1 and 3, and combinations 9-12 and 21-24 in Appendix F), the relationship between these two sights was exactly reversed. In the authors' opinion the reversal was undesirable but did not warrant the discarding of options. Finally, segment generalizability ranks were summed for each combination. The sum of the ranks

*The six gunner nonprecision objectives which were required by the stratification plan were drawn by selecting two items from each of three segments. The first was drawn in terms of segment generalizability. Since these three were night engagements, the highest ranking preferred day engagement was chosen as the second item to obtain a day/night balance.

ranged between 15 and 21. The five alternatives having the lowest rank sums (i.e., 15) were selected as the set from which to choose one combination.

Two candidates (i.e., combinations 3 and 4 in Appendix F) were eliminated from further contention because of their inclusion of more night than day objectives.* The three remaining candidates (i.e., combinations 13, 14 and 16) differed only in terms of the objectives used to represent HEP/BEEHIVE and SABOT/HEAT precision engagements. One combination (16) involved firing three engagements by primary methods during the day and one backup engagement at night. It was eliminated in favor of the remaining two candidates, both firing the backup engagement during the day. The final candidates (13 and 14) differed only in their HEP/BEEHIVE precision objectives. The choice was a moving target HEP engagement at night and a daylight BEEHIVE engagement, or a day/night reversal of these two objectives. Combination 13 was finally selected in which HEP was fired at a moving target during the day and BEEHIVE was fired at an area target at night.

As indicated earlier, selection of the tank commander/main gun objectives was straightforward. Choice of machinegun items was also relatively straightforward. For each machinegun objective where an option was available, the day objective was selected. To achieve a day/night balance among the .50 caliber engagements, one night objective (320) was replaced by the next highest ranking day engagement (298) belonging to the same segment.

MODEL TABLE VIII

The 28 engagements comprising the model Table VIII are presented in Tables 5 and 6. Main gun exercises are listed in the former and machinegun items appear in the latter. To provide a frame of reference in determining how representative the model exercises are of the gunnery domain, they were compared to the exercises comprising three other tables: one version of a USAARMS Table VIII presented in a draft of the revised FM 17-12; a USAREUR Table VIII described in USAREUR Reg. 350-704; and

*This step was reasonable because of the relatively large number of night, machinegun exercises which were eventually drawn. When combined, the two components provided a more nearly even split between day and night objectives.

Table 5. TABLE VIII MAIN GUN ENGAGEMENTS*

	Job Objective No.	Crew Member	Weapon	Mode of Firing	Firing Vehicle Motion	Target Motion	Target Type	Visibility	Day or Night	Fire Control Instrument	Ammu-nition	Target Range	Primary or Backup Method
DAYTIME													
1.	(244)	G	MG	BS	MH	M	TNK/LAV	VIS	D/N	TEL	SH	<1600	B
2.	(3)	G	MG	BS	M	S	TNK/LAV	VIS	D/N	GPD	SH	<1600	P
3.	(258)	TC	MG	P	MH	M	TNK/LAV	VIS	D/N	RFD	SH	500-4400	B
4.	(40)	G	MG	P	MH	M	TNK/LAV	VIS	D/N	TEL	SH	500-3200	B
5.	(35)	G	MG	P	MH	S	TNK/LAV	VIS	D/N	GPD	SH	500-4400	P
6.	(70)	TC	MG	P	MH	S	BKR/CRW	VIS	D/N	RFD	HEP	500-4400	B
7.	(262)	G	MG	P	MH	M	TSV	VIS	D/N	TEL	HEP	500-1600	P
NIGHTTIME													
1.	(83)	G	MG	RCLD	S	M	TNK/LAV	VAL	N	TEL	SH	<4400	B
2.	(89)**	G	MG	RCLD	S	S	TRPS	VAL	N	GPI	BEE	<1000	P
3.	(239)	TC	MG	RCLD	S	M	TNK/LAV	VAL	N	RFD	SH	<4400	B
4.	(10)	G	MG	BS	S	S	TNK/LAV	VAL	N	GPI	SH	<1000	P
5.	(63)***	G	MG	P	MH	S	TRPS	VAL	N	GPD	BEE	500-4400	P
6.	(27)	TC	MG	BS	M	S	TNK/LAV	VAL	N	RFD	SH	<1600	B
Exercises that may be substituted for BEEHIVE engagements.													
**	(274)	G	MG	RCLD	S	M	TSV/CRW	VAL	N	GPI	HEP	<1000	P
***	(56)	G	MG	P	MH	S	BKR/CRW	VAL	N	TEL	HEP	500-3200	P

* See Table 2 for descriptions of engagement conditions.

Table 6. TABLE VIII MACHINEGUN ENGAGEMENTS*

	Job Objective No.	Crew Member	Weapon	Mode of Firing	Firing Vehicle Motion	Target Motion	Target Type	Visibility	Day or Night	Fire Control Instrument	Ammu-nition	Target Range	Primary or Backup Method
DAYTIME													
1.	(133)	G	CX	NP	M	S	TRPS	VIS	D/N	IS	7.62	< 900	P
2.	(281)	TC	CX	NP	MH	S	TSV	VIS	D/N	RFD	7.62	< 900	B
3.	(119)	G	CX	NP	M	M	TSV/CRW	VIS	D/N	IS	7.62	< 900	P
4.	(307)	TC	.50	NP	MH	S	AIR	VIS	D/N	TPD	.50	< 2300	P
5.	(105)	G	CX	NP	M	S	TSV	VIS	D/N	IS	7.62	< 900	P
6.	(298)	TC	.50	NP	M	S	LAV/CRW	VIS	D/N	TPD	.50	< 1600	P
NIGHTTIME													
1.	(206)	G	CX	RCLD	S	M	TSV/CRW	VAL	N	GPI	7.62	< 900	P
2.	(214)	TC	CX	RCLD	S	S	TSV	VAL	N	RFI	7.62	< 900	B
3.	(282)	G	CX	NP	MH	S	TSV	VAL	N	IS	7.62	< 900	B
4.	(319)	TC	.50	NP	MH	S	AIR	VAL	N	TPI	.50	< 1000	P
5.	(139)	G	CX	NP	M	S	TRPS	VAL	N	GPI	7.62	< 900	P
6.	(296)	TC	CX	NP	MH	M	TSV/CRW	VAL	N	RFI	7.62	< 900	B
7.	(294)	G	CX	NP	MH	M	TSV/CRW	VAL	N	GPI	7.62	< 900	P
8.	(313)**	TC	.50	NP	MH	M	AIR	VAL	N	TPI	.50	< 1000	P
9.	(316)	TC	.50	NP	S	S	AIR	VAL	N	TPD	.50	< 2300	P
Exercises that may be substituted for moving aircraft target engagement.													
**	(230)	TC	.50	NP	MH	M	LAV	VAL	N	TPI	.50	< 1000	P

* See Table 2 for descriptions of engagement conditions.

the set of prototype exercises mentioned earlier in this report which were obtained by sampling all items strictly in terms of generalizability. Both the USAARMS and USAREUR tables were simply chosen as exemplars of alternative gunnery tests.

The relative coverage of engagement conditions provided by these four tables is shown in Table 7. With respect to the USAARMS and USAREUR tables the model Table VIII contains more exercises, the bulk of these consisting of coax engagements. The model table also includes more tank commander, range-card-lay, moving-to-a-halt, and moving target engagements. Further, the various fire control instruments are all represented and are reasonably balanced. The balancing of conditions is particularly noticeable when the day/night and fire control instrument frequencies of the prototype Table VIII are compared to those in the model Table VIII. These results suggest that the change in sampling procedure had the desired effects.

In Tables 5 and 6 alternative objectives (i.e., 274, 56, and 230) were listed for use in the event that BEEHIVE ammunition could not be fired and moving aircraft targets were unavailable. The substitution of these objectives modifies the conditions represented by the model Table VIII as listed in Table 7. A static target, two troop targets, an aircraft target, and a gunner's daylight periscope engagement are lost. A moving target, TSV/CRW, BKR/CRW, and LAV targets, and a telescope engagement are gained.

A more incisive basis for comparison is in terms of behavioral element coverage. The more comprehensive table will include a greater number of elements. Table 8 compares the four tests by listing those elements, from among the total set of 112, which are not contained within a given test. An "X" within a given column means that the Table VIII in question has not provided for coverage of a specific element. The code numbers of missing elements are given along the left margin of Table 8 together with the frequencies with which they occur in the gunnery domain. Elements not listed in the first column are covered in all of the tables. Descriptions of the elements appear in Appendix C.

Table 7. FREQUENCY OF ENGAGEMENT CONDITIONS REPRESENTED IN FOUR DIFFERENT TABLE VIII's

CONDITION		USAARMS	USAREUR	Prototype	Model
Crew Member:	Gunner	16	13	16	16
	Tank Commander	6	3	12	12
Weapon:	Main Gun	12	11	13	13
	Coax	6	2	10	10
	.50 Caliber	4	3	5	5
Firing Mode:	Battle Sight	8	5	4	4
	Precision	3	4	6	6
	Range Card Lay	1	2	5	5
	Non Precision	10	5	13	13
Firing Vehicle Motion:	Stationary	15	4	7	7
	Moving	5	9	6	7
	Moving to a Halt	2	3	15	14
Target Motion:	Stationary	16	13	17	17
	Moving	6	3	11	11
Target Type:	Tank/Light Armored Vehicle	10	11	9	9
	Thin Skinned Vehicle	2	3	5	5
	Troops	6	2	4	4
	Light Armored Vehicle/ Crew Served Weapon	2	–	–	1
	Aircraft	2	–	5	4
	Thin Skinned Vehicle/ Crew Served Weapon	–	–	5	5
Fire Control Instrument:	Gunner's Daylight Periscope	7	9	1	3
	Gunner's Infrared Periscope	2	1	7	5
	Rangefinder	2	–	4	5
	Rangefinder-Metascope	–	–	3	2
	Telescope	2	1	6	4
	Infinity Sight	5	2	2	4
	Tank Commander's Daylight Periscope	3	3	3	3
	Tank Commander's Infrared Periscope	1	–	2	2
Day/Night:	Day	11	10	7	13
	Night	11	6	21	15
	Total	22	16	28	28

Table 8. BEHAVIORAL ELEMENTS* NOT INCLUDED IN FOUR DIFFERENT TABLE VIII'S.

	Element No.	Domain Frequency	USAARMS	USAREUR	Prototype	Model
TANK COMMANDER ELEMENTS	5	17		X		
	9	12	X	X		
	10	1	X	X	X	X
	11	8	X			
	16	4	X	X		
	27	4		X	X	X
	28	43		X		
	29	5	X	X		
	30	2		X		
	31	41	X	X		
	32	8		X		
	34	5	X	X	X	X
	36	67		X		
	37	33		X		
	38	4		X		
	39	67		X		
	40	37		X		
	41	30	X	X		
	43	20	X	X		
	44	17			X	
	45	9			X	X
	46	22	X	X		
	47	8	X	X	X	X
	48	17		X		
	49	42		X		
	51	18	X	X		
	52	3	X	X	X	X
	53	15		X		
	54	3	X		X	X
	Percent Missing		14/52= 26.9%	25/52= 48.1%	8/52= 15.4%	7/52= 13.5%
GUNNER ELEMENTS	58	24		X		
	59	12	X	X		
	60	23	X	X		
	63	67		X		
	69	8	X			
	70	4	X	X		
	71	5		X		
	72	4	X	X	X	X
	74	27				X
	75	5	X	X	X	X
	77	5	X			
	80	18	X	X		
	81	20	X			
	82	15	X	X		
	90	18		X		
	91	12	X			
	94	7		X		
	Percent Missing		11/40 27.5%	12/40= 30.0%	2/40= 5.0%	3/40= 7.5%
DRIVER ELEMENTS	Percent Missing		0/6= 0%	0/6= 0%	0/6= 0%	0/6= 0%
LOADER ELEMENTS	104	16	X	X		
	105	27		X		
	110	18		X		
	111	9	X	X		
	112	16	X	X		
	113	27		X		
	Percent Missing		3/14= 21.4%	6/14= 42.9%	0/14= 0.0%	0/14= 0.0%
Total Percent Missing:			28/112 25.0%	43/112= 38.4%	10/112= 8.9%	10/112= 8.9%

* See Appendix C for description of elements.

The model Table VIII provides for coverage of over 90% of the elements in the gunnery domain. Of the ten which are missed, only one (74) occurs relatively frequently in the domain (27 times). This element ("gunner lays the rangeline on the center of target vulnerability"), is associated with precision engagements in which the telescope is used to fire at stationary targets. It does not appear in the model Table VIII because the various precision telescope engagements are fired at moving targets. The nine remaining missing elements are relatively rare, occurring from one to nine times in the domain of 266 objectives.

When the alternative objectives are substituted in the model Table VIII, the number of elements not covered rises to 16 or 14.3% of the total set. Substitution of the machinegun engagement (i.e., 230 for 313) does not affect element coverage. Substitution of the two main objectives (i.e., 274 and 56 for 89 and 63) results in the loss of seven elements (i.e., 9, 16, 59, 70, 104, 111, 112; see Appendix C), all of which occur relatively infrequently and are uniquely associated with BEEHIVE engagements. The main gun substitutions provide for coverage of one additional element (i.e., 74).

Comparisons between the model Table VIII and the USAARMS and USAREUR tables simply indicate that different combinations of objectives provide for the coverage of different elements. Consequently, when choosing one table over another one must be aware of the tradeoffs in element coverage which may be involved. The relatively smaller number of elements contained in the USAARMS and USAREUR Table VIII's is probably due to use of different (or implicitly constrained) gunnery domains and/or emphasis on sampling criteria other than generalizability.

The evaluations of condition and element coverage supported selection of the model objectives for a test of crew proficiency in tank gunnery. Finalizing the model Table VIII required decisions about ammunition allocations and specific tank-to-target ranges.

Ammunition. The main gun portion of the domain was defined in terms of the conditions and behaviors associated with first round engagements. A question that arises, therefore, is whether main gun model Table VIII exercises should be fired with one or two rounds of ammunition. Subject to the availability of resources, two rounds per engagement are recommended.

It will be recalled that subsequent-round engagements, after a first-round miss, were not used to define the domain because of: 1) the impossibility of specifying a priori which method of subsequent engagement would be used; and 2) the high similarity of behavioral elements involved in firing initially and adjusting fire.* Although subsequent-round engagements were not considered in defining the domain of objectives, there is no reason to eliminate them from the test of gunnery proficiency. Adjustment of fire is a natural sequel to the main gun objectives and has been formally included as one component of the standards for crew proficiency contained in FM 17-12 (1977).

Target Range. As part of the description of each gunnery objective an envelope of tank-to-target ranges has been specified. In each case the ranges are those which are permissible, given current doctrine, and within the capabilities of the M60A1AOS system. In order to finalize the model Table VIII it was necessary to choose specific ranges with which to represent the various engagements. As will be discussed in the section on implementation, such specifications (and adherence to them) are required in order to provide testing conditions which are standard for all crews.

The seven battlesight engagements were accompanied by range envelopes extending out to 1100 meters for HEAT ammunition and out to 1600 meters for SABOT. The objectives were arbitrarily split into HEAT and SABOT engagements and for each type the permissible ranges were represented. Assignment of a specific range to a particular objective was done randomly.

---

*Emerging gunnery doctrine of the future suggests that subsequent rounds will be fired by reinitiating the engagement rather than, for instance, making a burst-on-target adjustment.

Similar assignment of ranges was carried out for the six precision and three range-card-lay-to-direct-fire exercises. The upper limit on possible ranges for these objectives was assumed to exceed the capabilities of existing range facilities. Consequently, the upper range was arbitrarily fixed at a distance which presumably can be implemented (i.e., 2500 meters). The lower limit was set so that some of these engagements overlapped with battlesight exercises. Within these limits a representative sample of ranges was chosen and assigned at random to objectives.

The same approach was used to specify machinegun-to-target ranges. Because of the overlap among several of the envelopes and their narrower boundaries, many exercises were assigned essentially equivalent ranges. Whenever an envelope was encountered which contained extended ranges, the greater tank-to-target distances were chosen.

Summary. The model Table VIII tank qualification course is presented in Table 9 for main gun engagements and in Table 10 for machinegun engagements. Each exercise is described in terms of the target to be neutralized, the method of engagement, range to target, fire control instrument and type of illumination (for night engagements). The applicable crew standards are also given as specified in FM 17-12 (1977).

The remaining sections of this report focus exclusively on the model Table VIII. The discussion centers on scoring procedures to be used as well as considerations during implementation of the test.

Table 9. TABLE VIIIA, B TANK QUALIFICATION COURSE: Part I. Main Gun

Unit ____________ Tank Crew ____________ TC ____________

Gunner ____________ Loader ____________ Driver ____________

	EXERCISE		CONDITIONS	NO. OF ROUNDS	AMMUNITION	STANDARD
VIII A	1	MOVING TANK	Battlesight 1600 m Moving to a halt Telescope	2	TPDS-T	Engage in 5 seconds. Hit in 10 seconds.
	2	TANK FRONT SHOT	Battlesight 1000m On the move Gunner's periscope	2	HEAT-TP-T	Engage in 5 seconds. Hit in 10 seconds.
	3	MOVING TANK (TC)	Precision 1700 m Moving to a halt Rangefinder	2	TPDS-T	Engage in 10 seconds. Hit in 15 seconds.
	4	MOVING TANK	Precision 1700 m Moving to a halt Telescope	2	HEAT-TP-T	Engage in 10 seconds. Hit in 15 seconds.
	5	TANK SILHOUETTE	Precision 2000 m Moving to a halt Gunner's periscope	2	TPDS-T	Engage in 10 seconds. Hit in 15 seconds.
	6	BUNKER/ CREW WEAPON (TC)	Precision 2200 m Moving to a halt Rangefinder	2	HEP-TP-T	Engage in 10 seconds. Hit in 15 seconds.
	7	MOVING TRUCK	Precision 1200 m Moving to a halt Telescope	2	HEP-TP-T	Engage in 10 seconds. Hit in 15 seconds.
VIII B	1	MOVING TANK	Range card lay to direct fire 1900 m Stationary vehicle Telescope, flare	2	HEAT-TP-T	Hit within 5 minutes of reaching referenced position.
	2	TROOPS	Range card lay to direct fire 900 m Stationary vehicle Gunner Periscope, infrared	2	APERS	Hit within 5 minutes of reaching referenced position.
	3	MOVING TANK (TC)	Range card lay to direct fire 1400 m Stationary vehicle Rangefinder, flare	2	TPDS-T	Hit within 5 minutes of reaching referenced position.
	4	TANK FRONT SHOT	Battlesight 800 m Stationary vehicle Gunner's Periscope, infrared	2	HEAT-TP-T	Engage in 10 seconds. Hit within 15 seconds.
	5	TROOPS	Precision 1700 m Moving to a halt Gunner's periscope, flare	2	APERS	Engage in 15 seconds. Hit within 20 seconds.
	6	TANK FRONT SHOT (TC)	Battlesight 1300 m On the move Rangefinder, flare	2	TPDS-T	Engage in 10 seconds Hit within 15 seconds.

NOTES

1. During conduct of the table, target acquisition time, time to first-round hit, and time to second-round hit (if needed) are recorded. Scoring is then accomplished using a variety of procedures.
2. Emphasis is on achieving a target hit in the shortest possible time. Bonus points are given for ammunition conservation, and second round is not fired if the first round hits.
3. Crew duties are NOT scored on Table VIII for purposes of qualification
4. Three main gun rounds have been allocated for warm-up and zero confirmation (two rounds for day, one round for night). The least expensive round (HEP-TP-T) should be used for warm-up purposes and the highest muzzle velocity ammunition (TPDS-T) should be used for zeroing.
5. As an alternative, the second VIIIB engagement may be fired at a moving truck with HEP.
6. As an alternative, the fifth VIIIB engagement may be fired at a bunker with HEP and the telescope.
7. "Engage in 5 seconds" refers to the time from the alert element of the initial fire command or laying of the main gun for direction (whichever occurs earlier) to the firing of the first round. A second round, if needed, must be fired within 5 seconds of a first round miss.
8. Flare illumination may be replaced with white light illumination from another tank. Daylight standards should then be used

Table 10. TABLE VIII A, B TANK QUALIFICATION COURSE: Part II. Machinegun

		EXERCISE	CONDITIONS	NUMBER ROUNDS	STANDARD
VIIIA	**1**	**TROOPS**	300 m On the move Infinity sight	100 Coax	Engage within 5 seconds. Obtain 3/5s coverage.
	2	**TRUCK**	900 m Moving to a halt Rangefinder	50 Coax	Engage within 5 seconds. Obtain 1 tracer hit.
	3	**MOVING TRUCK**	700 m On the move Infinity sight	50 Coax	Engage within 5 seconds. Obtain 1 tracer hit.
	4	**AIRCRAFT**	2200 m Moving to a halt Tank Commander's periscope	100 Cal .50	Engage within 5 seconds. Obtain 1 tracer hit.
	5	**TRUCK**	500 m On the move Infinity sight	50 Coax	Engage within 5 seconds. Obtain 1 tracer hit.
	6	**TROOP CARRIER**	1500 m On the move Tank Commander's periscope	100 Cal .50	Engage within 5 seconds. Obtain 3/5s coverage.
VIIIB	**1**	**MOVING TRUCK**	300 m Stationary vehicle Gunner's periscope, infrared, RCLD	50 Coax	Engage within 10 seconds. Obtain 1 tracer hit.
	2	**TRUCK**	500 m Stationary vehicle Metascope, infrared, RCLD	50 Coax	Engage within 10 seconds. Obtain 1 tracer hit.
	3	**TRUCK**	900 m Moving to a halt Infinity sight, flare	50 Coax	Engage within 10 seconds. Obtain 1 tracer hit.
	4	**AIRCRAFT**	900 m Moving to a halt Tank Commander's periscope, infrared	100 Cal .50	Engage within 10 seconds. Obtain 1 tracer hit.
	5	**TROOPS**	700 m On the move Gunner's periscope, infrared	100 Coax	Engage within 10 seconds. Obtain 3/5s coverage.
	6	**MOVING TRUCK**	300 m Moving to a halt Metascope, infrared	50 Coax	Engage within 10 seconds. Obtain 1 tracer hit.
	7	**MOVING TRUCK**	500 m Moving to a halt Gunner's periscope, infrared	50 Coax	Engage within 10 seconds. Obtain 1 tracer hit.
	8	**MOVING AIRCRAFT**	900 m Moving to a halt Tank Commander's periscope, infrared	100 Cal .50	Engage within 10 seconds. Obtain 1 tracer hit.
	9	**AIRCRAFT**	2000 m Stationary vehicle Tank Commander's periscope, flare	100 Cal .50	Engage within 10 seconds. Obtain 1 tracer hit.

NOTES 1. During conduct of the table, engagement and hit times are recorded. Scoring is than accomplished using a variety of procedures.
2. As an alternative to VIII B, Exercise 8, a light armored vehicle may be engaged.
3. Flare illumination may be replaced with white light illumination from another tank. Daylight standards should then be used.
4. Metascope engagements are fired by the tank commander.

TABLE VIII SCORING PROCEDURES

Scoring issues and procedures are discussed in this section as a function of the four principal uses to which Table VIII results may be put. The section begins with a discussion of the kinds of data which are to be collected during the administration of the model table. Application of these data to specific uses is then treated separately for each potential scoring purpose.

PERFORMANCE DATA

The model Table VIII consists of 28 engagements, in which targets are to be neutralized with the coax, .50 caliber, and/or 105 mm. weapons. In each of these engagements data will be obtained that indicate a target hit or miss, the time required to conduct the engagement, and the quality of crew performance as measured by selected process variables.

Hit/Miss Data. The primary measure for each engagement will be an indication of whether the indicated target is hit or not. In machine-gun engagements of point targets a hit will be recorded when a minimum of one tracer round strikes the target (FM 17-12, 1977). Hits will be recorded for area and suppressive fire engagements when 3/5s coverage of arrayed targets is obtained (USAREUR Reg. 350-704, 1976). In main gun engagements a hit will be scored when the round strikes or passes through the target. A second round will not be fired unless the first round misses. In either event, which round strikes the target will be recorded. Following guidelines indicated in FM 17-12, credit for a hit will not be given if the (wrong) ammunition fired is incapable of destroying the target.

Engagement Time. Engagement time is defined as the time required for a tank crew to accomplish the single integrated act of acquiring, engaging, and hitting or destroying a threat target. When measuring speed of crew engagements, however, data are needed on several of the underlying components of behavior. Modifications are recommended, therefore, to the procedure outlined in FM 17-12 which results in continuous timing and recording of one overall time for the entire engagement.

In keeping with the crew standards specified for the M60A1AOS tank and an interest in diagnostic information, time data are needed for each of the following components: 1) acquisition time (i.e., the time from target appearance, as indicated by any of the alternative definitions in FM 17-12, to the alert element of the initial fire command or laying of the main gun for direction, whichever occurs earlier); 2) time to fire first round (i.e., the time from target acquisition as signified above to firing of the first round); 3) time to fire subsequent round (i.e., time from the first firing to the second firing); and 4) time to achieve a hit (i.e., the time from target acquisition to either a first or second round hit). In addition to these measures, range-card-lay-to-direct-fire (RCLDF) engagements will include the time from reaching the referenced firing position to securing target hits.

Collection of these data will place fairly heavy demands on scoring personnel who should be thoroughly briefed and trained in data collection procedures. The component measures would be obtained by permitting a stopwatch to run (from target appearance as defined in FM 17-12, or upon reaching a referenced firing position in RCLDF exercises) while recording elapsed time to salient events in seconds.

Process Variables. In addition to the primary data described above, information would be obtained on a number of ancillary aspects of crew performance. When multiple targets were presented within the context of a single engagement, the order in which the targets were engaged would be recorded. Similarly, in machinegun engagements the crew's technique of fire delivery would be qualitatively evaluated. In each engagement qualitative evaluations would be made of crew duties and interactions among crew positions (e.g., communication protocols). Finally, as suggested in FM 17-12, crews would be scored qualitatively on their use of terrain and quantitatively on their conservation of ammunition. These data and the speed and accuracy information would be used to evaluate crews for the purposes described below.

CREW QUALIFICATION

Basic hit/miss and component speed-of-engagement data should be used to develop scores which indicate whether or not crews are qualified in tank gunnery. The first step in this process is to determine whether crew performance on a given engagement is up to specified standards, scoring each engagement on a pass/fail basis. The second step is to aggregate the pass/fail information on individual engagements (i.e., the crew did or did not perform to standards) into a decision about crew qualification with respect to the domain of gunnery as a whole.

Crew Standards. The crew standards specified in FM 17-12, and presented earlier in this report, currently serve as the basis for all training and subsequent evaluation. However, a major caution is urged in applying these standards, or those proposed by others. As Kraemer, Boldovici, and Boycan (1975) suggest,

> Gunnery standards, and standards for all combat performance, should not be set on the basis of expert judgment, for if the experts are wrong, our gunners will be in trouble 'when the flag drops.' Nor should standards be set on the basis of the normative performance of our own trainees or qualified gunners. Normative data can tell us how good we are, but not how good we need to be. Standards for combat performance should be set on the basis of the best available information about the enemy's capability. Knowing that our gunners can meet arbitrarily established opening and closing time standards of 5 and 7 seconds provides little comfort if the enemy can open in 4 seconds and close in 6. Information about enemy gunnery capabilities must be made available to guide development of training and job performance standards (p. 42).

This caveat requires that gunnery standards, including those in FM 17-12 which have been adopted in the current effort, be subjected to continual scrutiny and revised as changes in our own or the enemy's doctrine and equipment warrant.

The crew standards presented earlier provide a basis for determining whether performance is adequate, that is, whether a crew should be passed or failed on a given engagement. Most of the standards, in fact, imply the application of multiple criteria during the evaluation of a given objective. For example, during a daylight battlesight engagement the crew must not only open fire within five seconds, but must also fire a subsequent round within five seconds (if needed) and obtain a target hit within ten seconds. The failure to satisfy any one of these standards, therefore, amounts to inadequate performance of the overall objective.* In an analogous sense, armor crewmen during the TCGST must demonstrate their ability to assemble and replace the breechblock and to accomplish this within six minutes. No credit is received if the breechblock is assembled but not replaced, or if it is assembled and replaced in seven instead of six minutes.

Score Aggregation. Application of the crew standards to relevant hit and time data will generate pass and fail (or 1 and 0) scores for each engagement. The next step is to aggregate these 28 individual statements about the adequacy of crew performance into one summary statement about the crew's performance on the domain as a whole. The aggregate or summary score would reflect the number of model Table VIII engagements passed (or failed) and would be indicative of the proportion of items the crew would succeed on had they been tested on the entire domain of gunnery objectives. For example, if a crew equalled or exceeded the specified performance standards on 21 out of 28 items, their performance in the aggregate would be characterized as succeeding on 75% of the engagements--both in the model table and in the larger domain.

However, a theoretically complex psychometric problem underlies the aggregation of component scores. At issue is whether the individual items or engagements measure the same construct, permitting a pooling or aggregation which is logically meaningful. In the extreme, were no form

*Occasionally, complaints are heard that this procedure is unduly stringent. In such cases, however, the argument is rarely with the scoring approach itself. Instead, the various standards come under attack. In those cases in which the standards have been arbitrarily set, this criticism often turns out to be beneficial, resulting in needed refinements and more realistic specification of what serves to distinguish acceptable from unacceptable performance.

of aggregation defensible, performance would have to be considered on an item by item basis. Empirical procedures such as Rasch modeling (Wright, 1967, Wright & Panchapakesan, 1969; Wright, 1977, and Steinheiser and Epstein, in press) can be used to shed light on this issue by scaling test items. These procedures, however, cannot be applied until a prototype of the test has been used to generate data. They also require large amounts of data obtained from repeated tests of the same individuals (or crews). Consequently, development of aggregation procedures for the model Table VIII proceeded on rational grounds.

Two alternatives suggested themselves. The first assumes that the domain which has been defined represents a single construct--tank gunnery. It further assumes that items drawn from the domain and used to comprise the model Table VIII can, by definition, be pooled to provide an estimate of performance in the domain at large based on a single aggregate score. The USAARMS Table VIII, although scoring performance differently, shares this view and aggregates scores across engagements. The second approach assumes that the domain can be divided into at least two components of gunnery, one representing the main gun and the other representing the machineguns. In this approach, a score would be aggregated to represent each component, and crew competence would be evaluated in terms of each. This latter approach is similar in many respects to that used in the USAREUR Table VIII. Part scores are in essence calculated and evaluated against a level of competence specified separately for machinegun engagements (e.g., obtain hits on all of them) and main gun engagements (e.g., obtain 7 out of 11 hits, etc.).

The second approach has been adopted for use with the model Table VIII since the underlying cluster analyses clearly identified two distinct super-clusters--main gun and machinegun. Accordingly, the performance of each crew must be characterized in terms of the proportion of the 13 main gun engagements which are conducted satisfactorily as well as in terms of the proportion of 15 machinegun exercises which are performed to standard. To qualify, a crew must demonstrate acceptable levels of proficiency on both of these two major components of the domain.

Score Interpretation. We have indicated how speed and accuracy data on any given engagement are compared to specified standards of performance to determine whether the crew has succeeded on that item. The question which now arises is how the two aggregate scores, which represent the pooling of item successes and failures, are to be interpreted. What does a score of 80% of main gun engagements performed satisfactorily mean? What criterion is used to evaluate such performance? In seeking an answer to these questions, it is useful to keep a number of criterion-referenced testing concepts in mind. The brief treatment given below is paraphrased along lines developed extensively by Shaycoft (in press) and others (Kriewall, 1972; Millman, 1973; and Popham and Husek, 1969).

For each tank crew firing the model Table VIII there are two scores (i.e., the percentages of main gun and machinegun test items passed) and two levels of competence (i.e., the percentages of main gun and machinegun engagements in the domain which the crew can perform to standards). The test scores are used as estimates of level of competence in the domain, based on at least an assumed monotonic relation between the two (e.g., the higher the crew's test score, the more items in the domain they are believed to be capable of handling). Paralleling the notions of score and level of competence are the concepts of standard of competence and cutting score. The standard of competence is a standard of proficiency which is arbitrarily set for a specified domain. In main gun tank gunnery, for example, it is that degree of mastery of the main gun portion of the domain which has been agreed upon as signifying crew qualification. For example, the standard of competence for main gun qualification might be set at 100%, 90%, or at any other level of competence in the domain. The cutting score is used operationally and is expressed as the percentage of Table VIII main gun items which must be passed if one is to infer that the standard of competence has been reached (in the domain).

Crew qualification should be based, therefore, on a comparison of the crew's Table VIII main gun and machinegun aggregate scores and the established cutting scores. When the crew's score (e.g., 85%) equals or exceeds the cutting score (e.g., 80%), the interpretation is that the crew's level of competence equals or exceeds the standard of competence which has been specified. Should their score (e.g., 85%) fall below the cutting score (e.g., 90%), their level of competence would be interpreted as falling below the indicated standard. As indicated earlier, both the main gun and the machinegun cutting scores must be equalled or exceeded in order for a crew to be deemed qualified.

Setting Standards and Cutting Scores. One of the most important steps in implementing the scoring approach described above is setting the standard of competence and cutting score. Unfortunately, guidelines to aid in this process are few and far between. As Shaycoft (in press) points out,

> The standard of competence to be set is not a function of the test itself but of the nature of the subject-matter area covered [the domain] and the purpose for which the test is being used (p. 71).

Cognizant armor personnel must, therefore, come to grips with crew qualification in terms of its implications for proficiency across the entire domain of gunnery. For example, should the 70% criterion used in the USAARMS Table VIII be formally adopted in the model Table VIII? Can one afford to count on crews, some of whom will hypothetically not be able to deal with 30% of the engagements they undertake? Alternatively, is a standard of 100% too stringent? These issues bear directly on the choice of a cutting score which, if equalled or exceeded, will support the inference that the standard of competence has been met. Here again, a number of options are possible including setting the cutting score higher than, lower than, or equal to the standard of competence. Resolution of these issues is complex and requires consideration of the effects of test length, level of competence, standard of competence, and cutting score on the accuracy of qualification decisions.

Mistakes will occasionally be made when judging crews as qualified or unqualified on the basis of their Table VIII performance. Such mistakes are known as classification errors. They refer to the fact that in some instances, because we are only dealing with a small sample of the domain, a crew may exceed the cutting score on the test by chance even though their true level of competence or functioning in the domain is below the standard. Such crews, who inadvertently qualify, are known as "false positives." Alternatively, crews may perform below the test cutting score by chance, when in fact, their true level of competence in the domain exceeds the standard. These crews, who inadvertently fail to qualify, are known as "false negatives." As several authors have indicated (i.e., Millman, 1973; Novick, & Lewis, 1974; Macready, Epstein, Steinheiser, & Mirabella, 1976; and Steinheiser, & Snyder, 1976) the binomial model can be used as an analytic aid to examine the tradeoffs between test length, cutting score, level of crew competence and probable rate of misclassification.

Test length has been fixed in the model Table VIII, 13 items supporting the evaluation of main gun performance, and 15 items being used to assess machinegun performance. At issue, therefore, are the misclassification rates which result when different cutting scores are used. These rates will vary as a function of the true levels of competence of each of the crews being evaluated. Consider the data presented in Table 11 which are the probabilities of misclassification error derived from the binomial model. Each column is defined by a crew's true level of competence with respect to the gunnery domain. Each row represents a test of given length, ranging from five to 28, and an aggregate score (i.e., the number of items correct) which would be needed to pass that test were the cutoff set at a given level. Four cut-off values are shown ranging from 95% at the top of the table to 70% at the bottom of Table 11.

The datum in any cell formed by the intersection of a column and a row is the percentage of crews who would be misclassified. For example, consider the data presented under the 95% cutting score section (Part I) of Table 11. Entries to the left of the vertical line represent the probabilities that crews, who although less competent than the 95% level, would pass more than 95% of the items on the test, and thus would be incorrectly judged as being at or above the 95% level of competence. They

Table 11. Probability of Misclassification Error

	Test Length	Passing Score	True Level of Crew Competence: False Positives 50%	60%	70%	80%	90%	False Negatives 95%
I. Cutting Score: Pass <u>95%</u> of Test Items	28	27	–	–	–	.02	.22	.41
	25	24	–	–	–	.03	.27	.36
	20	19	–	–	.01	.07	.39	.26
	15	14	–	.01	.04	.17	.55	.17
	13	12	–	.01	.06	.23	.62	.13
	10	9	.01	.05	.15	.38	.74	.08
	5	5	.03	.08	.17	.33	.59	.08
II. Cutting Score: Pass <u>90%</u> of Test Items	28	25	–	–	.02	.16	.31	.05
	25	22	–	–	.03	.23	.24	.03
	20	18	–	–	.04	.21	.32	.08
	15	14	–	.01	.04	.17	.45	.17
	13	12	.01	.01	.06	.23	.37	.14
	10	9	.01	.05	.15	.38	.26	.09
	5	5	.03	.08	.17	.33	.41	.23
III. Cutting Score: Pass <u>80%</u> of Test Items	28	22	–	.03	.22	.32	.02	–
	25	20	–	.03	.19	.38	.03	–
	20	16	.01	.05	.24	.37	.04	–
	15	12	.02	.09	.30	.35	.06	.01
	13	10	.05	.17	.42	.25	.03	–
	10	8	.06	.17	.38	.32	.07	.01
	5	4	.19	.34	.53	.26	.08	.02
IV. Cutting Score: Pass <u>70%</u> of Test Items	28	20	.02	.15	.47	.09	.02	–
	25	18	.02	.15	.49	.11	–	–
	20	14	.06	.25	.39	.09	–	–
	15	10	.15	.40	.28	.06	–	–
	13	9	.13	.35	.35	.10	.01	–
	10	7	.17	.38	.35	.11	.01	–
	5	4	.19	.34	.47	.26	.08	.02

are false positives. Entries to the right of the vertical lines are the probabilities that crews whose true levels of competence are 95% or greater would pass fewer than 95% of the test items, and thus would be incorrectly judged as being below the 95% level of competence.

From an examination of the rows of Table 11 corresponding to the test lengths of each part of the model Table VIII (i.e., 13 for the main gun portion and 15 for the machinegun portion), it is clear that there is an uncomfortably high probability that crews will be misclassified, especially when their true level of competence lies just above, just below, or precisely at the cutting score. For example, assume that the standard of competence was specified as 95%, and the cutting score set correspondingly at 95%. On the 13-item main gun portion of the model Table VIII, fully 13% of the crews whose true level of competence was precisely at the standard would fail to score highly enough to be deemed qualified; on the other hand, 62% of the crews whose true level of competence was five percentage points below the cutting score would score sufficiently high to be deemed qualified. These errors occur strictly by chance, and it is impossible to identify in any real-life situation a specific crew as a false positive or false negative. While expected misclassification rates decline rapidly as the true level of competence moves farther from the cutting score, if many crews in the population have true competences near the established standard of competence, as would be reasonable to suspect, the overall rate of misclassification is likely to be unacceptably high.

There are two ways within this kind of pass/fail scoring system to manipulate the parameters of the binomial model in order to improve the expected misclassification rates. The first is to increase the test length; it is generally the case that if appropriate cutting scores are used, the overall probability of misclassification decreases, as test length increases. Assuming that the model Table VIII as proposed herein is adopted, the only simple way of increasing the length is to administer the entire table more than once. If modifications of the present model table are to be considered, then the sampling fraction might be modified to produce a larger sample of test items. In either of these cases, the cost of administering the table increases.

Another alternative which might be employed to mitigate the problem is to adopt a different system of selecting cutting scores. Up to this point it has been generally assumed that the cutting score ought to be selected to reflect (i.e., be equal or nearly equal to) the standard of competence, and that a single cutting score will be employed to reflect a single standard of competence. However, one may work from a different direction entirely, that is, to begin with a consideration of the consequences of various kinds of misclassification, and then to derive various standards and cutting scores which minimize those consequences. As will shortly be seen, this approach leads to a "red/amber/green," or pass/marginal/fail scoring system rather than a pass/fail dichotomy. It will be seen that such a system can be used to alleviate certain kinds of misclassification problems, and also may be highly congruent with the Army's use of qualification information. This approach is developed in the next section, which also considers the consequences of adopting such a rationale.

Determining Crew Qualification. In describing this approach an arbitrary standard of competence must first be adopted, by which a crew could definitely be judged qualified, e.g., 95%.* Thus, for any crew whose true level of competence is 95% on the main gun portion of the domain and on the machinegun portion of the domain, it could be labelled "qualified" with high confidence. The consequences of mislabelling such a crew would seem to be high. For example, if a truly competent crew is labelled "un-

---

*Some will certainly argue that in a criterion-referenced test of this sort a more proper standard of competence would be 100%. Certain practical problems arise that make such a standard inappropriate and unfair. In particular, random round-to-round dispersion produced by the main gun weapon system will make perfect performance impossible to achieve in many circumstances, even for 100% competent crews. A companion report (Fingerman, 1977) discusses this problem in more detail, and explores methods of adjusting the testing situation to make relatively high levels of performance (e.g., 95%) attainable by highly competent crews. Without such adjustments, even the 95% standard proposed here is impractical.

78

qualified" and sent for remedial training, that crew is no longer available for an appropriate job role (e.g., front-line duty). This line of reasoning suggests that a cutting score which minimizes the probability that a truly qualified crew would be labelled as unqualified would be desirable. Examining the bottom portion of Table 11, a crew whose true level of competence is 95% or greater has less than a 1% chance of falling below a cutting score of 70%, i.e., of failing to pass at least 70% of the items on either a 13 or 15 item test; this probability is still about 1% for a cutting score of 80%. If the cutting score is to be set so that the probability of misclassifying (failing) a crew whose true level of competence is at least 95% (the standard of competence) is to be less than 1%, the appropriate cutting scores are 9 out of 13 (or approximately 70%) on the main gun portion of the table, and 11 out of 15 (or approximately 73%) on the machinegun portion according to the binomial model. In other words, fewer than 1% of the crews who are at or above 95% true competence will pass less than 70% of the main gun exercises, and fewer than 1% will pass less than 73% of the machinegun exercises. Since qualification demands that both portions be satisfactorily completed, a maximum of 2% of the truly qualified crews would be incorrectly labelled as "unqualified" if these cutting scores were adopted.

This same argument can be applied in a slightly inverted fashion, that is, one might specify a standard of "incompetence." In this case, a standard of 70% might be adopted; the implication is that crews whose true level of competence is at or below 70% can be confidently labelled "unqualified." Once again, consider the consequences of misclassification: the cost of incorrectly labelling a poor crew "qualified" could be quite high (c.f. "I would not want him protecting my flank.") This approach suggests selecting a cutting score which minimizes the probability that a poor crew would be labelled as qualified by the model table. Using the binomial model in this situation, fewer than 6% of crews whose true level of competence was 70% or below would pass more than 12 out of 13 items (approximately 92%) on the main gun portion of the model table, and fewer than 4% of these crews would pass the machinegun portion with a cutting score of 14 out of 15 (about 93%). In other words, if these

79

cutting scores were adopted, one could be at least 90% confident that crews whose true level of competence was less than 70% would not be labelled as qualified.

These lines of argument generate two cutting scores for each portion of the table, one designed to minimize the chance of labelling a truly qualified crew as unqualified, and the other designed to minimize the chance of labelling a truly poor crew as qualified. This leaves the slightly ambiguous position depicted in Figure 2. Crews whose performance places them in the upper third of the figure on both portions of the table can be labelled as qualified, knowing in fact that the chance is quite small that any such crew's true level of competence is less than or equal to 70%. Crews falling in the lower third on either portion of the model Table VIII can be labelled as unqualified, and in fact very few crews with a true competence of 95% or above will be so classified. Crews falling in the middle are, in some sense, in limbo. The central portion of the figure corresponds to the situation referred to in statistics as the region in which "judgment is suspended". In the simplest terms, one does not know enough about crews who perform in this region to confidently label them either "qualified" or "unqualified." The lack of discriminability for these crews derives from the problem with which this section started: for tests of lengths such as have been specified for the model Table VIII, there are always high probabilities of misclassification error for some true levels of competence. While the three-category approach will allow many crews to be unambiguously categorized, one must remain uncertain about the rest.

In this situation there are again at least two alternatives. First, the crews who fall in the middle might be retested, effectively doubling the test length. They would then be classified based on their total performance, using a single cutting score with increased confidence. As in the preceding discussion of increasing test length, the cost implications may make this approach undesirable. A second approach would be to label the crews who fall in the middle "marginally qualified."

The use of three classifications (i.e., qualified, marginally qualified, unqualified) in connection with Table VIII would certainly not

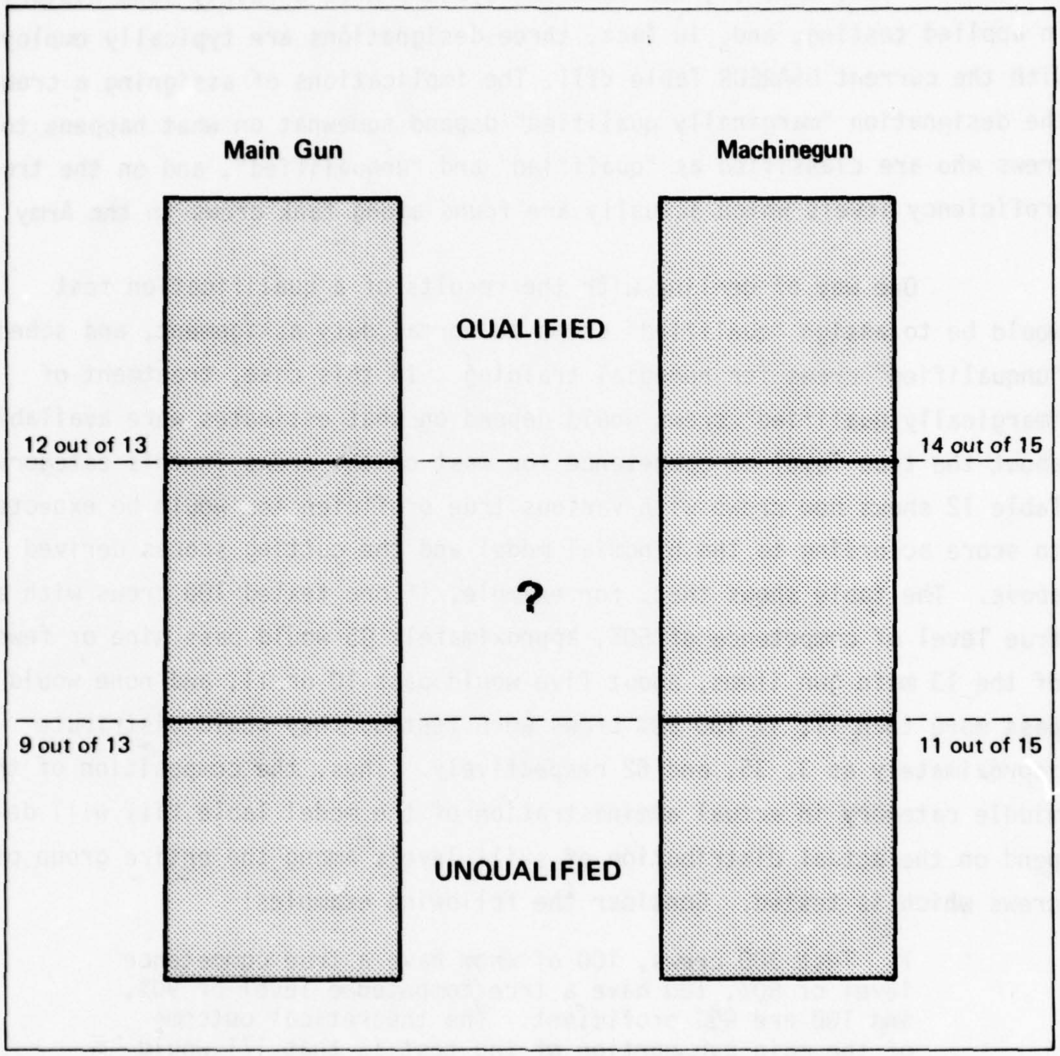

Figure 2. Multiple cutting scores.

be novel. "Red, amber, green" scoring systems have a fairly long history in applied testing, and, in fact, three designations are typically employed with the current USAREUR Table VIII. The implications of assigning a crew the designation "marginally qualified" depend somewhat on what happens to crews who are classified as "qualified" and "unqualified", and on the true proficiency levels which actually are found among tank crews in the Army.

One way of dealing with the results of a qualification test would be to assign "qualified" crews to normal duty assignment, and schedule "unqualified" crews for remedial training. In this case, treatment of "marginally qualified" crews would depend on what estimates were available about the true level of competence for most of the crews in this category. Table 12 shows how crews with various true proficiencies would be expected to score according to the binomial model and the cutting scores derived above. The table shows that, for example, if one tested 100 crews with a true level of competence of 50%, approximately 95 would pass nine or fewer of the 13 main gun items, about five would pass 10 or 11, and none would pass more than 11; if 100 90% crews were tested, they would distribute approximately as 3, 35, and 62 respectively. Thus, the composition of the middle category in actual administration of the model Table VIII will depend on the actual distribution of skill levels among the entire group of crews which is tested. Consider the following examples:

1. Test 300 crews, 100 of whom have a true competence level of 80%, 100 have a true competence level of 90%, and 100 are 95% proficient. The theoretical outcome of the main gun portion of the test is that 171 would pass, 28 would fail, and 101 would fall in the middle. The average true competence level of all crews tested is 88.3%, and the average competence level of the crews in the middle classification is 85.54.

2. Test 300 crews, 100 of whom have a true competence level of 50%, 100 have a competence level of 60%, 50 are at the 70% level, 30 at 80%, 10 at 90%, and 10 at 95%. The predicted outcome for this group of 300 crews is (approximately) 26 passes, 215 failures, and 59 crews in the middle. In this case, the average competence level of all the tested crews is 62.5%, and the average competence level of the crews in the middle category would be about 70%.

Table 12. Expected Distributions On Each Portion Of The Model Table

	True Level of Crew Competence	Score and Category: ≤ 9 Unqualified	10–11 Marginally Qualified	≥ 12 Qualified
I. Main Gun (13 items)	**50%**	**.95**	**.05**	**—**
	60%	**.83**	**.16**	**.01**
	70%	**.58**	**.36**	**.06**
	80%	**.25**	**.52**	**.23**
	90%	**.03**	**.35**	**.62**
	95%	**—**	**.14**	**.86**
	True Level of Crew Competence	≤ 11 Unqualified	12–13 Marginally Qualified	≥ 14 Qualified
II. Machinegun (15 items)	**50%**	**.98**	**.02**	**—**
	60%	**.91**	**.08**	**.01**
	70%	**.70**	**.26**	**.04**
	80%	**.35**	**.48**	**.17**
	90%	**.06**	**.39**	**.55**
	95%	**.01**	**.16**	**.83**

The first example characterizes what would happen were the model Table VIII administered to a population of tank crews which was fairly proficient, while the second illustrates the outcome for a poorer group of crews. The decision, for example, to schedule "marginally qualified" crews for extensive remedial training as opposed to specific remedial training on certain, often-failed, exercises (see the section on skill diagnosis, below) or for some additional practice might thus be based on an estimate of the true distribution of proficiency in the testee population. If the crews are believed to be fairly proficient, then most of those who fall into the middle classification on Table VIII can be assumed to be fairly proficient, and some practice will significantly improve their capabilities; if, on the other hand, one is concerned that the average level of competence is low, one might then conclude that crews labelled "marginally qualified" require additional training. Table VIII performance data summed over an entire unit, for example, might be useful in estimating the crew-population skill distribution.

Scoring Crew Qualification. With the three-category scheme for determining crew qualification, the assignment of a particular designation is straightforward. Table 13 has been prepared using the standards and cutting scores derived in the preceding section to show the outcome for various combinations of scores on the two portions of the Table VIII. Crews who receive 12 or more points on the main gun portion (i.e., perform 12 or more exercises to the specified standards) and 14 or more points on the machinegun portion, would be designated "qualified" crews. Crews which score nine or fewer points on the main gun portion or 11 or fewer points on the machinegun portion would be designated "unqualified". The remainder of the possible score combinations would be associated with the "marginally qualified" designation. In other words, crews which "pass" both portions of the test are qualified, crews which "fail" either portion are unqualified, and crews whose scores fall into the middle category on both portions of the test, or into the "pass" category on one portion and the middle on the other portion are marginally qualified.

The introduction to this chapter made the point that the raw data from Table VIII should be retained regardless of the particular scoring

Table 13. Determining Crew Qualification

Number of Machinegun Exercises Performed to Standard	Number of Main Gun Exercises Performed to Standard: 0 – 9	10 – 11	12 – 13
0 – 11	Unqualified	Unqualified	Unqualified
12 – 13	Unqualified	Marginally Qualified	Marginally Qualified
14 – 15	Unqualified	Marginally Qualified	Qualified

purposes and analyses which are actually to be addressed. The same is true of the "pass" and "fail" information on each portion of the test. While this information is only used as an intermediate step in scoring for crew qualification, it may be of direct use when other purposes and uses of Table VIII information are considered (see the remaining sections in this chapter).

One other point remains to be made with regard to scoring for crew qualification. The standards of competence (i.e., 95%) and "incompetence" (i.e., 70%) seem entirely justifiable to the staff of the present project; however, circumstances may dictate that other standards be adopted (c.f. footnote, p. 66). It should be noted that the binomial model may be reapplied to determine new cutting scores for each portion of Table VIII: the lower cutting score is selected such that crews who have a true level of competence which equals or exceeds the standard of competence should have less than a 1% chance of "failing" each part of the table; the upper cutting score is selected such that crews which have a true level of competence which equals or is less than the standard of "incompetence" should have less than a 5% chance of "passing" each part. Note also that, if the relative importance of false positives and/or false negatives changes in the future, the probabilities employed in deriving the cutting scores (i.e., 1% and 5%) may also be adjusted.

PREDICTION OF COMBAT EFFECTIVENESS

Underlying many military test and evaluation efforts is the (usually) implicit assumption that the obtained data can provide estimates of what may be expected in the way of performance in a combat environment. Often, however, too much is expected either in terms of the accuracy and validity of the predictions or the breadth and scope of the performance involved. Such is the case with tank crews. Demonstration of proficiency on Table VIII cannot be unequivocally translated into consequent effectiveness in combat. The caveats surrounding such predictions arise for a variety of reasons.

Issues in Estimating Combat Effectiveness. First, and foremost, there is the ubiquitous criterion problem. Predictions of a crew's effectiveness in combat can be readily made, but it is virtually impossible to substantiate or validate them since the criterion effectiveness information must be gathered in a combat environment. Nor is this mere hair splitting. The stresses, havoc, and chaos of actual combat presumably can exert profound influences, of either a positive or negative nature, on crew performance. If these pressures are absent or of a different nature when the criterion data are collected, there is no reason to believe that an accurate assessment has been or can be made.

A second problem concerns what aspects of tank crew performance are subsumed under a concept as potentially broad as combat effectiveness. Within the context of the current project, effectiveness presumably translates into the neutralization of targets which is accomplished within specified standards of speed and accuracy. Effectiveness within this context is essentially equivalent, therefore, to survival in combat -- to neutralizing the target before one's own tank is "blown away".

This raises the third problem. Whether a crew prevails in such an engagement is, as we have seen, a function of more than their proficiency in marksmanship. Coolness under fire and the ability to make a variety of sound tactical decisions will also strongly influence the outcome. It is evident, therefore, that proficiency in marksmanship is necessary but not sufficient for survival in combat. Having raised these caveats, it is fair to say that the performance of crews on the model Table VIII can, in a very limited sense, be used to predict their survival in combat.

Scoring. The key issue in attempting to make predictions of survival in combat lies in identifying those families within the gunnery domain which are particularly germane or of concern. Two criteria to use in pinpointing these families are their criticality to survival and their probability of occurrence. Cognizant armor personnel, considering engagement conditions which are likely to exist within their own local areas of operation, must apply the two criteria, selecting those items from the model Table VIII on which the performance of crews should be scrutinized, while discarding the rest. For example, one area commander might, because of the probable conditions in his area, and the missions which are most likely to be assigned to his unit, focus on those exercises in the model Table VIII which represent range-card-lay-to-direct-fire families. Another commander might discard those same families and focus instead on precision SABOT/HEAT engagements.

Once the relevant engagements were designated, scoring of these exercises would be identical to the approach described for crew qualification. Measures of crew speed and accuracy would be obtained and compared to the specified performance standards to determine success or failure on each engagement. These scores would then be aggregated over the set of combat-related engagements, and the proportion of combat-related exercises which were performed up to standard would be an index of combat effectiveness. If decisions about combat effectiveness were to be criterion-referenced, cutting scores would have to be selected using the same procedures described above for qualification scoring. Such cutoff points would take into account the impact of system unreliability on decision errors.

Crews who performed poorly or who failed to meet the cutting score specified for the set of critical engagements would be predicted to succumb in combat. The prediction would be based on their demonstrated lack of skills necessary for survival. Ironically, predictions of survival would not be possible for crews performing all of the critical exercises satisfactorily. At best one might indicate that they possessed some of the skills (i.e., marksmanship skills) particularly relevant to or necessary for survival, and would outlast crews who did not possess such skills. It is interesting to note that either outcome is independent of the crew qualification decision. A qualified crew might fail an exercise from a critical cluster while an unqualified crew (with respect to the domain as a whole) might be extraordinarily competent in one critical aspect of gunnery. Both crews might be slated for remediation but their training would focus on different regions of the domain.

SKILL DIAGNOSIS

Once crews have fired the model Table VIII, a number of questions can be asked about their performance that are independent of qualification decisions or predictions of survival in combat. For a given crew, for instance, one would like to know on what engagements (or families of engagements) the crew excels. Similarly, one would want to determine the kinds of exercises or the portions of the gunnery domain on which the crew experiences difficulty. In this latter case more detailed information would be useful, especially if it indicated the particular parameter of performance on which the crew was deficient (e.g., excessively long times to fire the first round) and possible causes for the deficiency (e.g., relatively long target acquisition times).

The model Table VIII is ideally suited to these diagnostic appraisals of performance. This is so because emphasis during administration of the test is placed on collection of raw crew performance data which are to be recorded and preserved to support a variety of analyses and decisions. As a consequence it is possible to refer back to a given engagement and retrieve basic facts such as, for example, "a hit was not secured," or "the time to fire the second round was three seconds." Other scores are derived and used for different purposes but the basic performance data are kept intact.

Armed with basic measures of performance, which have neither been transformed nor aggregated, it is possible to produce profiles of a crew's performance which provide for differential diagnosis of skill. By examining the peaks and valleys in such a profile the unit commander can, for instance, isolate specific facets of performance on which improvement is required (e.g., time to fire the first round) and the conditions under which practice is most likely to be beneficial for the crew (e.g., long range precision engagements).

Specification of remedial training is also possible for larger command units by describing unit performance in terms of the same profiles. When this is done, for instance, it may become apparent that all (or most) crews within a given battalion require excessive time to fire the first round of extended-range, precision engagements. When summed over sufficiently large numbers of crews, inputs of this type can be used to identify segments of the entire armor training program which may require modification. The ability to formulate specific remediation strategies which have payoff will depend on the aspects of gunnery performance that are monitored, and how the results of the diagnostic analysis are communicated to the parties concerned.

Performance Measures. In choosing a mix of measures which is to be monitored for diagnostic significance, one has to decide how much data he can afford to collect, specifically within the context of the model Table VIII. Some promising measures (e.g., gun camera sight pictures) may require expensive instrumentation while others (e.g., crew coordination) may require one or more "on-board" observers, etc.

In the model Table VIII a variety of outcome and process measures is readily available. Outcome measures, expressed in the form of raw data (e.g., "on engagement #6 the time to get off the first round was 17 seconds") or as deviations from a performance standard (e.g., "first round time on engagement #6 was seven seconds slower than permitted"), provide symptomatic information. On any given engagement one can establish which aspect of the outcome was satisfactory or unsatisfactory, and, relative to performance standards, determine how large a deviation existed. Outcome measures, therefore, can be used to pinpoint aspects of performance in need of remediation. Process measures, on the other hand, address cause and become of interest in potentially explaining why, for example, a given crew was inordinately slow on engagement #6. Although such measures can be addressed in their own right (e.g.,"communication problems between the driver and tank commander were evident in this crew"), they usually assume significance only when the outcome was unsatisfactory. In this case they may provide insights into the kind of remedial training which is needed.

In the model Table VIII, symptomatic information will be provided by the outcome measures specified for use in scoring qualification. These include hit-miss data, an indication of the round with which a hit is obtained, and information about the component times, specifically acquisition time, involved in an engagement. Several process measures might also be used. The first, concerning use of terrain, would be formally scored on each engagement by adopting the procedure currently used in the USAARMS Table VIII. A second measure, adopted from the USAREUR Table VIII would indicate the order in which multiple targets were engaged and whether that order was consistent with relative target threat. A third measure would focus on crew interaction, and be scored by comparing recorded communication protocols to appropriate doctrine on conduct-of-fire and crew coordination.

CREW MOTIVATION

In the past, interest in differentiating among crews in terms of their gunnery performance has led to the development of scoring systems which facilitated such discriminations. In the USAARMS Table VIII a point system is used which permits qualified crews (i.e., those scoring 1700 points or more) to be compared on relative grounds. Those crews scoring between 1900 and 2000 points (i.e., 90% or better) are designated "Distinguished" and awarded an "Expert" badge; those scoring 1800 to 1899 points (i.e., $< 90\%$, $\geq 80\%$) are termed "Superior" and given a "Sharpshooter" badge; and those scoring 1700 to 1799 points (i.e. $< 80\%$, $\geq 70\%$) are referred to as "Qualified" and given a "Marksman" badge. A two-category system is used to differentiate among crews who qualify on the USAREUR Table VIII. Crews are designated as "green" if, among other criteria, they achieve nine or more hits on 11 targets and "amber" if they get seven or eight hits.

Data obtained from the model Table VIII can also be used to categorize or otherwise differentially reward qualified crews. A scoring system for this purpose must satisfy two straightforward conditions. First, it obviously must order crews in a manner which is consistent with the basic crew-qualification decision. In other words, all qualified crews should score higher than marginally qualified crews who should score higher than unqualified crews. Second, the system must be sufficiently sensitive to distinguish among the performance of crews, all of whom have qualified.

There are many scoring systems which could satisfy the two constraints mentioned above. For example, the point scoring system currently associated with the USAARMS Table VIII might be adapted for this purpose. For both the main gun and machinegun engagements, points would be awarded using sliding time scales. Points would also be awarded for technique and effect of machinegun fire, for ammunition conservation, and for appropriate use of terrain. Simpler alternative scoring systems might also be considered which would be less demanding for the personnel who would administer and score the test, while still being consistent with the constraints. Among the simplest would be a system which awarded two points to any crew which passed all of the main gun engagements (i.e., 13 out of 13, where 12 out of 13 is required for qualification), and one point to any crew which passed all of the machinegun engagements (i.e., 15 out of 15, where 14 out of 15 is required for qualification). This system would award "bonus points" for performing better than is required for qualification, and would weight main gun engagements more heavily than machinegun exercises; the point scores would range from one to three, and would only be available to crews who had qualified. Other variants of this scheme are obvious.

Choice of one over another of these systems is largely an arbitrary matter, as is the specification of distinct levels of performance (i.e., number of points) which would be differentially recognized and rewarded.* The real issue is whether such discriminations are truly useful. In the criterion-referenced approach to crew qualification advocated

*"Arbitrary" is not meant to imply "trivial" or "inconsequential." The design assumptions underlying the model Table VIII simply provide no special guidance with regard to such decisions. The choice of an existing method over a simpler one must be referred to Armor personnel more familiar with the status of crew motivation and morale. The selection of specific point levels might be made by these same personnel, once some normative performance data on the model Table VIII have been accumulated.

in this report, it may actually be counterproductive to force a distinction among crews when all of them have met exceedingly high standards. An alternative motivational device and reward system might instead be keyed to the single fact of qualification. The point scoring system would then provide an informal but acceptable way of generating competition among crews and representing their relative standings.

IMPLEMENTATION OF MODEL TABLE VIII

The steps involved in developing a model Table VIII test of crew proficiency in tank gunnery have focused thus far on specification of test items and elaboration of multi-purpose scoring procedures. This presentation would be incomplete without also discussing some of the issues involved in implementation of the test. These considerations can be characterized in terms of three broad topics. The first concerns the layout of the testing facility needed to support the model Table VIII. The second involves the actual programming of the engagements which are to be fired, either as single targets or multiple-target arrays. The third and final topic addresses the collection of quality data within a standardized test format.

RANGE FACILITIES

Guidelines for preparing, organizing, and using tank gunnery ranges are contained in FM 17-12 (1977) and will not be repeated here. These prescriptions detail the physical layouts and range control procedures which must be implemented to satisfy safety requirements and to expedite processing of crews. Implementation of the model Table VIII involves a number of facility considerations in addition to these basic concerns.

Physical Constraints. The model table is comprised of a series of rather demanding engagements which tend to emphasize two conditions. First, many targets are to be acquired, engaged and neutralized at relatively extended ranges, not only with respect to the weapon used, but also in terms of the amount of actual terrain required. Second, numerous engagements involve the use of moving targets and/or moving tanks. These two factors, both singly and in combination, impose major burdens on the physical test facility, particularly when, as has been assumed thus far in the report, firing is to be conducted with live ammunition.

Although range facilities currently under development (e.g., at Fort Knox) seem adequate, other facilities on which variations of Table VIII are currently fired may not be able to support specific exercises included in the model table. The primary restriction would probably be range-to-target, a situation which led to imposition of the arbitrary ceiling of 2500 meters on model Table VIII engagements. If an existing gunnery range must be used, and that facility cannot support long-range engagements, a compromise must be reached. In this case, the range specified for the model engagement in question must be reduced to a shorter distance, which is still at or near the maximum for the facility. When several such changes must be made, the recommendation would be to retain a broad band of ranges, and to the extent possible, minimize the number of precision engagements fired at battlesight ranges. Such deviation from the ranges prescribed in the model Table VIII is not desirable and should be kept to a minimum. It does, however, represent a practical and acceptable solution when limited facilities are available.

The compromise necessitated by an inability to implement moving tank and/or moving target engagements is a much more complicated matter. One can't simply substitute a stationary-tank or stationary-target exercise. The difference in the two situations is that changes in range don't change the mix of behavioral elements underlying crew performance. Changes in tank or target motion do alter the nature and sequence of the behavioral steps. Modifications of this type are, therefore, much more serious since they fundamentally change the nature of the job objective(s) being evaluated.

The key to the solution of this problem lies in performing behavioral element trade-off analyses of the type illustrated in Table 8. Candidate substitutes are evaluated by determining which elements are lost and which are gained. Only on this basis is it really possible to determine what impact a change in exercises has, and consequently, whether the candidate substitute should be used. When the engagement in question is the sole representative of one of the 16 clusters of objectives, or when several engagements are problematic, changes in the generalizability index should also be considered. The detailed job-objective descriptions needed to conduct these trade-off analyses have been reproduced and made available to the U.S. Army Armor School (Boldovici, Boycan, Fingerman & Wheaton, 1977).

Pop-Up Targets. A second consideration in implementation of the model Table VIII is the desirability of using pop-up targets to the extent possible. An obvious advantage of such targets is that they can facilitate evaluation of the crew's target acquisition performance. Much of the ambiguity is removed concerning the actual moment of target appearance. It is reasonable to assume, therefore, that more accurate estimates of acquisition time will be possible.

There is a second and perhaps more important advantage associated with the use of these kinds of targets in the model Table VIII. It stems not from the pop-up but rather from the knock-down feature. Scoring of a target hit should be far less ambiguous than in the past when heavy reliance was placed upon the judgment of one or more observers. Increased accuracy in determining hits is extremely important in the model Table VIII. A premium is placed on crew qualification, and, in order to qualify, the crew must neutralize essentially all the targets with which it is presented. As a consequence, there is no margin for error in determining a target hit or miss.

Target Size. Traditionally, the size of many of the targets used in tests of crew marksmanship has approximated the area presented by a tank's turret when engaged frontally. Target panels based on this principle are routinely used which measure approximately 2.3 by 2.3 meters. For testing purposes one can question the use of such targets in engagements conducted at both close and long ranges. Larger targets may provide better estimates of crew marksmanship particularly for ranges in excess of 1,000-1,200 meters.

At issue is the notion of weapon system reliability. For instance, as mentioned in FM 17-12-2 (1977), dispersion effects are frequently noted, in which a round, for a variety of reasons, does not strike precisely where the weapon is aimed. This situation represents an unreasonable penalty with which crews must contend during qualification when shooting at relatively small targets at longer ranges. In spite of being layed precisely on target, a miss may occur through no fault of the crew. Again, given the stringent standards for qualification in the model Table VIII, it is important that the effects of such system limitations be minimized.

One way of doing so is to systematically regulate target size, as a function of several parameters (e.g., engagement type, range). This approach is feasible, although a number of logistical (e.g., need for new targets) and psychometric (e.g. misclassification rates) issues would have to be addressed. The psychometric issues are briefly explored in a companion report, along with a preliminary examination of weapon system limitations on crew performance (Fingerman, 1977).

TARGET ARRAYS

The gunnery performance which the model Table VIII has been designed to address, namely marksmanship, consists of one portion of a larger domain of behavior. The other components of the domain, involving a variety of situations and crew behaviors, consist of various forms of tactical decision making. One of the most important of these, and one which is not far removed from marksmanship per se, is the engagement of multiple targets.

When a multiple-target array is encountered the crew must first prioritize all targets with respect to threat. This establishes the sequence in which they will be engaged. The crew must then bring such targets under fire, engaging them either sequentially, or if circumstances are appropriate, simultaneously. The question is the degree to which such engagements can be practiced and evaluated within the context of the model Table VIII.

The answer involves one overriding consideration. The Table VIII described in this report has been expressly designed to provide estimates of crew proficiency in marksmanship. During implementation of the test, therefore, care must be taken not to jeopardize attainment of this goal. Accordingly, the programming of sequential or simultaneous multi-target exercises must be judged in terms of how disruptive they might be to the assessment of marksmanship skill. Analysis of the situation suggests that sequential, multi-target engagements can be fielded during both the day and at night as well as within or across weapon systems. Such engagements should be minimally disruptive. Taking the lead from the USAARMS and USAREUR Table VIII's, several acceptable multiple-target engagements might be constructed. For example, during the daylight portion of the model table, the following would be reasonable (c.f., Tables 9 and 10):

1) Combine main gun engagement #2 (gunner, HEAT, battlesight, on the move, against a frontal tank) with machinegun engagement #3 (gunner, coax, infinity sight, on the move against a moving truck). Note that, while the order of engagement is important to assess decision making and/or to add realism, the marksmanship performance for both exercises, which is of critical interest in the model table, would be assessed without regard to which target is taken under fire first.

2) Combine main gun engagement #5 (gunner, SABOT, precision, moving-to-a-halt, against a tank) with main gun engagement #2. As in the previous example, marksmanship performance would be scored independently of the order in which targets were engaged, but a critique might point out that the nearer tank should be engaged first.

The night portion of the table is also amenable to the use of sequential, multiple-target engagements. For example, the following might be considered:

1) Combine main gun engagement #4 (gunner, HEAT, battlesight, against a tank, from a stationary position using infrared) and machinegun engagement #1 (gunner, coax, stationary against a moving truck, infrared). The ambush-type scenario here should also be scored primarily for marksmanship, but the presentation of two targets of different threat allows examination of the tank commander's tactical decision.

2) Combine tank commander machinegun engagements #8 and #9. This multiple-target engagement would involve the tank commander engaging a nearby aircraft using his infrared periscope while moving to a halt, and then engaging a faraway aircraft with flare illumination. Marksmanship performance on the two engagements taken separately is again of first priority in scoring for the model table, but the exercises might gain significantly in realism if presented in combination.

The judicious use of such exercises should be beneficial in training crews and diagnosing tactically relevant performance, while not jeopardizing the validity of the obtained data for scoring marksmanship skill.

The inclusion of targets which are to be engaged simultaneously is open to more question because of the potentially greater disruption which such exercises may have on basic crew marksmanship tasks. In particular, when the tank commander is firing the .50 caliber machinegun and the gunner simultaneously is firing the main gun or coax, some tank commander behaviors associated with the gunner's engagement may be omitted (Miller & Hayes, 1976). It would seem prudent to provide for simultaneous engagements elsewhere, perhaps during conduct of tables concerned with section/platoon distribution of fire against numerically superior forces.

QUALITY CONTROL

While the foregoing are important considerations, this last issue is crucial. Steps must be taken to administer the table in a consistent and standardized fashion, ensuring that all crews, insofar as possible, have been evaluated in terms of the same conditions and standards of performance. This requirement has implications for site-to-site variations and for the personnel supporting the test at any given site.

Test Standardization. Currently, area commanders can modify Table VIII to suit local conditions. As a consequence, the crews qualify by demonstrating their ability to perform different tasks, each set of tasks being somewhat site-dependent. This variation in local implementation is not consistent with the purposes of the Table VIII described in this report. The model Table VIII assumes rigorously controlled conditions which are not changed because of site specific weather or maintenance problems.

If rain precluded engagement of one of the targets at 2300 meters because of reduced visibility, the exercise would not be conducted at 700 meters where the target would presumably be distinct. Such a change may, as already mentioned above, subtly change the nature of what is being measured. Similarly, when a crew experiences malfunctions, such as a gun jamming, they would be permitted to fire the engagement they would otherwise miss. Such retesting would take place even in those instances where *crew negligence* or action directly contributed to the malfunction. The model Table VIII is not addressing such things as preventive maintenance. The test is one of how well the crew can shoot when told what to shoot at and how the target is to be engaged.

Implementation of this last requirement may be the single most difficult aspect of putting the model table into practice. In order to sample the gunnery domain adequately, crews would demonstrate proficiency using different methods of engagement. For example, the tank commander would fire some main gun engagements on his own; the gunner would use back-up methods, etc. The trick is to ensure that such engagements are actually performed by crews. How conformance with the requirements for each exercise can be guaranteed is not readily apparent. The most straightforward approach is to supply observers who certify that each engagement is conducted in strict accordance with the specified scenario (i.e., the tank commander actually does fire the round, or the gunner actually does use the specified backup sight). Crews would not fire freely, without control of engagement method, since the resulting variation would seriously weaken the overall test strategy, which is the sampling of specific conditions and behavioral elements. The liberal use of test examiners/observers therefore seems necessary.

Standardization of Procedure. The second step in assuring quality control is to train test personnel to record speed measures, determine target hit or miss, and to evaluate use of terrain, according to standardized, systematic procedures.

Formal training and checkout of test administrators are essential if the obtained data are to be of sufficient quality to support the qualification, motivation, and diagnostic functions of the model Table VIII.

Crew training in gunnery proceeds sequentially and is predicated on a building-block approach. Individual crew members first demonstrate their ability to perform a number of tasks related to their specific crew positions. Having demonstrated their individual proficiencies, the next step is to weld the crew members into a highly skilled and proficient team. One of the first opportunities for the team to demonstrate its prowess is in performance of representative tasks comprising the domain of gunnery marksmanship.

The present report has described the need for, and presented the steps involved in, developing a test instrument for use in assessing crew gunnery proficiency. The test which has emerged measures one facet of gunnery and does so within the context of a broader training and evaluation program. That is, once crews have demonstrated their proficiency in basic marksmanship, attention can turn to increasingly more sophisticated kinds of performance. These advanced skills, which involve additional coordination within and among crews, are all fundamentally predicated on the assumption that the crew can basically use its weapons systems to neutralize targets.

The next phase of research in this project continues to focus on the assessment of crew proficiency in gunnery. However, the context in which such assessment occurs is changed dramatically. Subsequent research and development will determine the degree to which model Table VIII exercises of the kind recommended in this report can be used on a simulated basis to obtain valid estimates of crew proficiency. Simulation-based testing, as an alternative to live-fire testing of Table VIII will, if feasible, permit reallocation of scarce ammunition and range resources to support development of more advanced tactical crew and unit skills.

REFERENCES

Boldovici, J.A., Wheaton, G.R., & Boycan, G.G. Selecting items for a tank gunnery test, Interim Report. Arlington, VA.: U.S. Army Research Institute for the Behavioral and Social Sciences, 1976.

Boldovici, J.A., Boycan, G.G., Fingerman, P.W., & Wheaton, G.R. Tank gunnery data handbook. Alexandria, VA.: U.S. Army Research Institute for the Behavioral and Social Sciences, 1977.

Dixon, W.J. (Ed.) BMDP: Biomedical computer programs. Berkeley, CA.: University of California Press, 1975.

Fingerman, P.W. A preliminary investigation of weapon-system dispersion and crew marksmanship. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, 1977.

Hartigan, J.A. Direct clustering of a data matrix. Journal of the American Statistical Association, 1972, 67, 123-129.

Hartigan, J.A. Clustering algorithms. New York, N.Y.: John Wiley, 1975.

Kraemer, R.E., Boldovici, J.A., & Boycan, G.G. Job objectives for M60A1AOS tank gunnery compared to proposed training. Vol. 1: Development and results. Arlington, VA.: U.S. Army Research Institute for the Behavioral and Social Sciences, 1975.

Kriewall, T.E. Aspects and applications of criterion-referenced tests. Paper presented at American Educational Research Association Annual Meeting; Chicago, IL., 1972.

Macready, G.B., Epstein, K.I., Steinheiser, F.H., & Mirabella, A. Methods and models for criterion-referenced testing. Arlington, VA.: U.S. Army Research Institute for the Behavioral and Social Sciences, Technical Paper, 1976.

Miller, E.E., & Hayes, J.F. Analysis of tank crew performance requirements for multiple-target engagements. Arlington, VA.: U.S. Army Research Institute for the Behavioral and Social Sciences, 1976 (Draft).

REFERENCES (Cont'd.)

Millman, J. Passing scores and test lengths for domain-referenced tests. Review of Educational Research, 1973, 43, 205-215.

Novick, M. & Lewis, C. Prescribing test length for criterion-referenced measurement. In C.W. Harris, M.C. Alkin, & W.J. Popham (Eds), Center for the Study of Evaluation Monograph Series in Evaluation III: Problems in criterion-referenced measurement. Los Angeles, CA.: Center for the Study of Evaluation at U.C.L.A., 1974.

Popham, W.J., & Husek, T.R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969 6, 1-9.

Shaycoft, M.E. Guide to the development, evaluation, and use of criterion-referenced tests. Palo Alto, CA.: American Institutes for Research, in press.

Smode, A.F., Gruber, A., & Ely, H. The measurement of advanced flight vehicle crew proficiency in synthetic ground environments. Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratories, 1962.

Steinheiser, F.H., Jr., and Epstein, K.I. An experimental investigation of the military police firearms qualification course. Alexandria, VA.: U.S. Army Research Institute, in press.

Steinheiser, F.H., Jr., & Snyder, C.W., Jr. Score quality issues related to individual and weapon crew criterion-referenced performance tests. Paper presented at the Military Testing Association Conference, Gulf Shores, AL., 1976.

U.S. Army Armor School (USAARMS). Tank gunnery. Field Manual (FM-17-12), Washington, D.C.: Headquarters, Department of the Army, in press, 1977.

U.S. Army Armor School (USAARMS). Tank gunnery for M60, M60A1, M60A1(AOS), and M48A5 tanks. Field Manual (FM-17-12-2), Washington, D.C.: Headquarters, Department of the Army, 1977.

U.S. Army, Europe (USAREUR). Training: Tank/sheridan/anti-tank crew qualification. Draft USAREUR Regulation 350-704. Heidelberg, Germany: Headquarters, USAREUR, May, 1976.

Wheaton, G.R. Scoring procedures for a tank gunnery test, Interim Report. Arlington, VA.: U.S. Army Research Institute for the Behavioral and Social Sciences, 1977.

Wright, B.D. Sample - free test calibration and person measurement. Paper presented at the Invitational Conference on Testing Problems, 1967.

Wright, B.D. Solving measurement problems with the Rasch model. Journal of Educational Measurement, in press, 1977, 14.

Wright, B.D. & Panchapakesan, N. A procedure for sample-free item analysis. Educational and Psychological Measurement, 1969, 29, 23-48.

APPENDIX A DOMAIN OF GUNNERY JOB OBJECTIVES

The following pages describe each of the 266 job objectives in terms of the conditions under which each engagement occurs. They are boxed into 16 groups as indicated by the cluster solution. Abbreviations used for various conditions appear below:

Crew member:
TC = tank commander

Weapon:
MG = main gun   COAX = coaxial machinegun (7.62 mm)
CAL.50 = .50 caliber cupola machinegun

Firing mode:
BS = battlesight   P = precision   NP = non-precision
RC = range card   RCLD = range card lay to direct fire

Firing vehicle and target motion:
S = stationary   M = moving   MH = moving to a halt

Target type:
TSV = thin-skinned vehicle   TNK = tank or heavy armored vehicle
LAV = light armored vehicle   BKR = bunker
CREW = crew-served weapon   AIR = aircraft
ALL = all targets

Target visibility:
VIS = visible   VAL = visible with artificial illumination
NVIS = not visible

Day or night:
D/N = day or night   N = night

Fire control instrument:
GPD = gunner's daylight periscope   GPI = gunner's infrared periscope
TEL = gunner's telescope   INF = gunner's infinity sight
RFD = rangefinder, daylight   RFI = rangefinder, infrared (metascope)
TPD = tank commander's daylight periscope   AUX = auxiliary fire controls
TPI = tank commander's infrared periscope

Ammunition:
SAB/HEAT = Sabot or Heat   COAX = 7.62 mm   CAL.50 = .50 caliber

Cluster No. 1

OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
1	GUNNER	MG	BS	S	S	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	<1600
10	GUNNER	MG	BS	S	S	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
9	GUNNER	MG	BS	S	S	TNK/LAV	VAL	N	GPD	SAB/HEAT	<1600
2	GUNNER	MG	BS	S	S	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	<1600
11	GUNNER	MG	BS	S	S	TNK/LAV	VAL	N	TEL	SAB/HEAT	<1600
5	GUNNER	MG	BS	S	M	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	<1600
16	GUNNER	MG	BS	S	M	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
15	GUNNER	MG	BS	S	M	TNK/LAV	VAL	N	GPD	SAB/HEAT	<1600
6	GUNNER	MG	BS	S	M	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	<1600
17	GUNNER	MG	BS	S	M	TNK/LAV	VAL	N	TEL	SAB/HEAT	<1600
242	GUNNER	MG	BS	MH	S	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	<1600
247	GUNNER	MG	BS	MH	S	TNK/LAV	VAL	N	TEL	SAB/HEAT	<1600
241	GUNNER	MG	BS	MH	S	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	<1600
3	GUNNER	MG	BS	M	S	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	<1600
13	GUNNER	MG	BS	M	S	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
12	GUNNER	MG	BS	M	S	TNK/LAV	VAL	N	GPD	SAB/HEAT	<1600
4	GUNNER	MG	BS	M	S	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	<1600
14	GUNNER	MG	BS	M	S	TNK/LAV	VAL	N	TEL	SAB/HEAT	<1600
7	GUNNER	MG	BS	M	M	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	<1600
19	GUNNER	MG	BS	M	M	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
18	GUNNER	MG	BS	M	M	TNK/LAV	VAL	N	GPD	SAB/HEAT	<1600
8	GUNNER	MG	BS	M	M	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	<1600
20	GUNNER	MG	BS	M	M	TNK/LAV	VAL	N	TEL	SAB/HEAT	<1600
249	GUNNER	MG	BS	MH	S	TNK/LAV	VAL	N	GPD	SAB/HEAT	<1600
248	GUNNER	MG	BS	MH	S	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
245	GUNNER	MG	BS	MH	M	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	<1600
253	GUNNER	MG	BS	MH	M	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
254	GUNNER	MG	BS	MH	M	TNK/LAV	VAL	N	GPD	SAB/HEAT	<1600
244	GUNNER	MG	BS	MH	M	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	<1600
252	GUNNER	MG	BS	MH	M	TNK/LAV	VAL	N	TEL	SAB/HEAT	<1600

Cluster No. 2

21	TC	MG	BS	S	S	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	<1600
25	TC	MG	BS	S	S	TNK/LAV	VAL	N	RFD	SAB/HEAT	<1600
23	TC	MG	BS	S	M	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	<1600
29	TC	MG	BS	S	M	TNK/LAV	VAL	N	RFD	SAB/HEAT	<1600
243	TC	MG	BS	MH	S	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	<1600
246	TC	MG	BS	MH	M	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	<1600
251	TC	MG	BS	MH	S	TNK/LAV	VAL	N	RFD	SAB/HEAT	<1600
256	TC	MG	BS	MH	M	TNK/LAV	VAL	N	RFD	SAB/HEAT	<1600
22	TC	MG	BS	M	S	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	<1600
24	TC	MG	BS	M	M	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	<1600
27	TC	MG	BS	M	S	TNK/LAV	VAL	N	RFD	SAB/HEAT	<1600
31	TC	MG	BS	M	M	TNK/LAV	VAL	N	RFD	SAB/HEAT	<1600

Cluster No. 3

OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
33	GUNNER	MG	P	S	S	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	5-4400
41	GUNNER	MG	P	S	S	TNK/LAV	VAL	N	GPD	SAB/HEAT	5-4400
35	GUNNER	MG	P	MH	S	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	5-4400
43	GUNNER	MG	P	MH	S	TNK/LAV	VAL	N	GPD	SAB/HEAT	5-4400
34	GUNNER	MG	P	S	S	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	5-3200
42	GUNNER	MG	P	S	S	TNK/LAV	VAL	N	TEL	SAB/HEAT	5-3200
36	GUNNER	MG	P	MH	S	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	5-3200
44	GUNNER	MG	P	MH	S	TNK/LAV	VAL	N	TEL	SAB/HEAT	5-3200
40	GUNNER	MG	P	MH	M	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	5-3200
48	GUNNER	MG	P	MH	M	TNK/LAV	VAL	N	TEL	SAB/HEAT	5-3200
37	GUNNER	MG	P	S	M	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	5-4400
45	GUNNER	MG	P	S	M	TNK/LAV	VAL	N	GPD	SAB/HEAT	5-4400
39	GUNNER	MG	P	MH	M	TNK/LAV	VIS	D/N	GPD	SAB/HEAT	5-4400
47	GUNNER	MG	P	MH	M	TNK/LAV	VAL	N	GPD	SAB/HEAT	5-4400
38	GUNNER	MG	P	S	M	TNK/LAV	VIS	D/N	TEL	SAB/HEAT	5-3200
46	GUNNER	MG	P	S	M	TNK/LAV	VAL	N	TEL	SAB/HEAT	5-3200

Cluster No. 4

OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
65	TC	MG	P	S	S	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	5-4400
66	TC	MG	P	MH	S	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	5-4400
67	TC	MG	P	S	S	TNK/LAV	VAL	N	RFD	SAB/HEAT	5-4400
68	TC	MG	P	MH	S	TNK/LAV	VAL	N	RFD	SAB/HEAT	5-4400
257	TC	MG	P	S	M	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	5-4400
259	TC	MG	P	S	M	TNK/LAV	VAL	N	RFD	SAB/HEAT	5-4400
258	TC	MG	P	MH	M	TNK/LAV	VIS	D/N	RFD	SAB/HEAT	5-4400
260	TC	MG	P	MH	M	TNK/LAV	VAL	N	RFD	SAB/HEAT	5-4400

Cluster No. 5

CREW MEMBER	OBJECTIVE NUMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
49	GUNNER	MG	P	S	S	BKR/CREW	VIS	D/N	GPD	HEP	5-4400
50	GUNNER	MG	P	S	S	BKR/CREW	VIS	D/N	TEL	HEP	5-3200
51	GUNNER	MG	P	MH	S	BKR/CREW	VIS	D/N	GPD	HEP	5-4400
52	GUNNER	MG	P	MH	S	BKR/CREW	VIS	D/N	TEL	HEP	5-3200
53	GUNNER	MG	P	S	S	BKR/CREW	VAL	N	GPD	HEP	5-4400
54	GUNNER	MG	P	S	S	BKR/CREW	VAL	N	TEL	HEP	5-3200
55	GUNNER	MG	P	MH	S	BKR/CREW	VAL	N	GPD	HEP	5-4400
56	GUNNER	MG	P	MH	S	BKR/CREW	VAL	N	TEL	HEP	5-3200
57	GUNNER	MG	P	S	S	TROOPS	VIS	D/N	GPD	BEEHIVE	5-4400
58	GUNNER	MG	P	S	S	TROOPS	VIS	D/N	TEL	BEEHIVE	5-1200
59	GUNNER	MG	P	MH	S	TROOPS	VIS	D/N	GPD	BEEHIVE	5-4400
60	GUNNER	MG	P	MH	S	TROOPS	VIS	D/N	TEL	BEEHIVE	5-1200
61	GUNNER	MG	P	S	S	TROOPS	VAL	N	GPD	BEEHIVE	5-4400
62	GUNNER	MG	P	S	S	TROOPS	VAL	N	TEL	BEEHIVE	5-1200
63	GUNNER	MG	P	MH	S	TROOPS	VAL	N	GPD	BEEHIVE	5-4400
64	GUNNER	MG	P	MH	S	TROOPS	VAL	N	TEL	BEEHIVE	5-1200
261	GUNNER	MG	P	S	M	TSV	VIS	D/N	TEL	HEP	5-1600
262	GUNNER	MG	P	MH	M	TSV	VIS	D/N	TEL	HEP	5-1600
263	GUNNER	MG	P	S	M	TSV	VAL	N	TEL	HEP	5-1600
264	GUNNER	MG	P	MH	M	TSV	VAL	N	TEL	HEP	5-1600
265	GUNNER	MG	P	S	M	TSV	VIS	D/N	GPD	HEP	5-1600
266	GUNNER	MG	P	MH	M	TSV	VIS	D/N	GPD	HEP	5-1600
267	GUNNER	MG	P	S	M	TSV	VAL	N	GPD	HEP	5-1600
268	GUNNER	MG	P	MH	M	TSV	VAL	N	GPD	HEP	5-1600

Cluster No. 6

71	TC	MG	P	S	S	BKR/CREW	VAL	N	RFD	HEP	5-4400
73	TC	MG	P	S	S	TROOPS	VIS	D/N	RFD	BEEHIVE	5-4400
74	TC	MG	P	MH	S	TROOPS	VIS	D/N	RFD	BEEHIVE	5-4400
75	TC	MG	P	S	S	TROOPS	VAL	N	RFD	BEEHIVE	5-4400
76	TC	MG	P	MH	S	TROOPS	VAL	N	RFD	BEEHIVE	5-4400
269	TC	MG	P	S	M	TSV	VIS	D/N	RFD	HEP	5-1600
270	TC	MG	P	MH	M	TSV	VIS	D/N	RFD	HEP	5-1600
271	TC	MG	P	S	M	TSV	VAL	N	RFD	HEP	5-1600
272	TC	MG	P	MH	M	TSV	VAL	N	RFD	HEP	5-1600
69	TC	MG	P	S	S	BKR/CREW	VIS	D/N	RFD	HEP	5-4400
70	TC	MG	P	MH	S	BKR/CREW	VIS	D/N	RFD	HEP	5-4400
72	TC	MG	P	MH	S	BKR/CREW	VAL	N	RFD	HEP	5-4400

Cluster	OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
Cluster No. 7	77	GUNNER	MG	RC	S	S	ALL	NVIS	D/N	AUX	HEP	<4400
	84	GUNNER	MG	RCLD	S	S	BKR/CREW	VAL	N	GPD	HEP	<4400
	85	GUNNER	MG	RCLD	S	S	BKR/CREW	VAL	N	GPI	HEP	<1000
	86	GUNNER	MG	RCLD	S	S	BKR/CREW	VAL	N	TEL	HEP	<3200
	88	GUNNER	MG	RCLD	S	S	TROOPS	VAL	N	GPD	BEEHIVE	<4400
	89	GUNNER	MG	RCLD	S	S	TROOPS	VAL	N	GPI	BEEHIVE	<1000
	90	GUNNER	MG	RCLD	S	S	TROOPS	VAL	N	TEL	BEEHIVE	<3200
	273	GUNNER	MG	RCLD	S	M	TSV/CREW	VAL	N	TEL	HEP	<1600
	274	GUNNER	MG	RCLD	S	M	TSV/CREW	VAL	N	GPI	HEP	<1000
	275	GUNNER	MG	RCLD	S	M	TSV/CREW	VAL	N	GPD	HEP	<1600
Cluster No. 8	78	GUNNER	MG	RCLD	S	S	TNK/LAV	VAL	N	GPD	SAB/HEAT	<4400
	83	GUNNER	MG	RCLD	S	M	TNK/LAV	VAL	N	TEL	SAB/HEAT	<4400
	81	GUNNER	MG	RCLD	S	M	TNK/LAV	VAL	N	GPD	SAB/HEAT	<4400
	80	GUNNER	MG	RCLD	S	S	TNK/LAV	VAL	N	TEL	SAB/HEAT	<4400
	79	GUNNER	MG	RCLD	S	S	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
	82	GUNNER	MG	RCLD	S	M	TNK/LAV	VAL	N	GPI	SAB/HEAT	<1000
Cluster No. 9	92	TC	MG	RCLD	S	S	TNK/LAV	VAL	N	RFD	SAB/HEAT	<4400
	234	TC	MG	RCLD	S	M	TNK/LAV	VAL	N	RFD	SAB/HEAT	<4400
Cluster No. 10	94	TC	MG	RCLD	S	S	BKR/CREW	VAL	N	RFD	HEP	<4400
	97	TC	MG	RCLD	S	S	TROOPS	VAL	N	RFD	BEEHIVE	<4400
	277	TC	MG	RCLD	S	M	TSV/CREW	VAL	N	RFD	HEP	<1600
Cluster No. 11	131	GUNNER	COAX	NP	M	S	TROOPS	VIS	D/N	GPD	COAX	<900
	133	GUNNER	COAX	NP	M	S	TROOPS	VIS	D/N	INF	COAX	<900
	132	GUNNER	COAX	NP	M	S	TROOPS	VIS	D/N	TEL	COAX	<900
	139	GUNNER	COAX	NP	M	S	TROOPS	VAL	N	GPI	COAX	<900
	138	GUNNER	COAX	NP	M	S	TROOPS	VAL	N	GPD	COAX	<900
	141	GUNNER	COAX	NP	M	S	TROOPS	VAL	N	INF	COAX	<900
	140	GUNNER	COAX	NP	M	S	TROOPS	VAL	N	TEL	COAX	<900
	129	GUNNER	COAX	NP	S	S	TROOPS	VIS	D/N	TEL	COAX	<900
	136	GUNNER	COAX	NP	S	S	TROOPS	VAL	N	TEL	COAX	<900
	128	GUNNER	COAX	NP	S	S	TROOPS	VIS	D/N	GPD	COAX	<900
	135	GUNNER	COAX	NP	S	S	TROOPS	VAL	N	GPI	COAX	<900
	134	GUNNER	COAX	NP	S	S	TROOPS	VAL	N	GPD	COAX	<900
	130	GUNNER	COAX	NP	S	S	TROOPS	VIS	D/N	INF	COAX	<900
	137	GUNNER	COAX	NP	S	S	TROOPS	VAL	N	INF	COAX	<900
	104	GUNNER	COAX	NP	M	S	TSV	VIS	D/N	TEL	COAX	<900

continued on next page

Cluster No. 11 (continued)

OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
112	GUNNER	COAX	NP	M	S	TSV	VAL	N	TEL	COAX	<900
103	GUNNER	COAX	NP	M	S	TSV	VIS	D/N	GPD	COAX	<900
111	GUNNER	COAX	NP	M	S	TSV	VAL	N	GPI	COAX	<900
110	GUNNER	COAX	NP	M	S	TSV	VAL	N	GPD	COAX	<900
105	GUNNER	COAX	NP	M	S	TSV	VIS	D/N	INF	COAX	<900
113	GUNNER	COAX	NP	M	S	TSV	VAL	N	INF	COAX	<900
288	GUNNER	COAX	NP	MH	M	TSV/CREW	VIS	D/N	INF	COAX	<900
292	GUNNER	COAX	NP	MH	M	TSV/CREW	VAL	N	INF	COAX	<900
290	GUNNER	COAX	NP	MH	M	TSV/CREW	VIS	D/N	GPD	COAX	<900
294	GUNNER	COAX	NP	MH	M	TSV/CREW	VAL	N	GPI	COAX	<900
295	GUNNER	COAX	NP	MH	M	TSV/CREW	VAL	N	GPD	COAX	<900
293	GUNNER	COAX	NP	MH	M	TSV/CREW	VAL	N	TEL	COAX	<900
289	GUNNER	COAX	NP	MH	M	TSV/CREW	VIS	D/N	TEL	COAX	<900
279	GUNNER	COAX	NP	MH	S	TSV	VIS	D/N	TEL	COAX	<900
283	GUNNER	COAX	NP	MH	S	TSV	VAL	N	TEL	COAX	<900
117	GUNNER	COAX	NP	M	M	TSV/CREW	VIS	D/N	GPD	COAX	<900
125	GUNNER	COAX	NP	M	M	TSV/CREW	VAL	N	GPI	COAX	<900
124	GUNNER	COAX	NP	M	M	TSV/CREW	VAL	N	GPD	COAX	<900
119	GUNNER	COAX	NP	M	M	TSV/CREW	VIS	D/N	INF	COAX	<900
127	GUNNER	COAX	NP	M	M	TSV/CREW	VAL	N	INF	COAX	<900
114	GUNNER	COAX	NP	S	M	TSV/CREW	VIS	D/N	GPD	COAX	<900
121	GUNNER	COAX	NP	S	M	TSV/CREW	VAL	N	GPI	COAX	<900
120	GUNNER	COAX	NP	S	M	TSV/CREW	VAL	N	GPD	COAX	<900
116	GUNNER	COAX	NP	S	M	TSV/CREW	VIS	D/N	INF	COAX	<900
123	GUNNER	COAX	NP	S	M	TSV/CREW	VAL	N	INF	COAX	<900
118	GUNNER	COAX	NP	M	M	TSV/CREW	VIS	D/N	TEL	COAX	<900
126	GUNNER	COAX	NP	M	M	TSV/CREW	VAL	N	TEL	COAX	<900
280	GUNNER	COAX	NP	MH	S	TSV	VIS	D/N	GPD	COAX	<900
284	GUNNER	COAX	NP	MH	S	TSV	VAL	N	GPI	COAX	<900
285	GUNNER	COAX	NP	MH	S	TSV	VAL	N	GPD	COAX	<900
278	GUNNER	COAX	NP	MH	S	TSV	VIS	D/N	INF	COAX	<900
282	GUNNER	COAX	NP	MH	S	TSV	VAL	N	INF	COAX	<900
100	GUNNER	COAX	NP	S	S	TSV	VIS	D/N	GPD	COAX	<900
107	GUNNER	COAX	NP	S	S	TSV	VAL	N	GPI	COAX	<900
106	GUNNER	COAX	NP	S	S	TSV	VAL	N	GPD	COAX	<900
102	GUNNER	COAX	NP	S	S	TSV	VIS	D/N	INF	COAX	<900
109	GUNNER	COAX	NP	S	S	TSV	VAL	N	INF	COAX	<900
101	GUNNER	COAX	NP	S	S	TSV	VIS	D/N	TEL	COAX	<900
108	GUNNER	COAX	NP	S	S	TSV	VAL	N	TEL	COAX	<900
115	GUNNER	COAX	NP	S	M	TSV/CREW	VIS	D/N	TEL	COAX	<900
122	GUNNER	COAX	NP	S	M	TSV/CREW	VAL	N	TEL	COAX	<900

	OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
Cluster No. 12	201	GUNNER	COAX	RCLD	S	S	TSV	VAL	N	GPD	COAX	<900
	202	GUNNER	COAX	RCLD	S	S	TSV	VAL	N	GPI	COAX	<900
	203	GUNNER	COAX	RCLD	S	S	TSV	VAL	N	TEL	COAX	<900
	204	GUNNER	COAX	RCLD	S	S	TSV	VAL	N	INF	COAX	<900
	208	GUNNER	COAX	RCLD	S	M	TSV/CREW	VAL	N	INF	COAX	<900
	209	GUNNER	COAX	RCLD	S	S	TROOPS	VAL	N	GPD	COAX	<900
	210	GUNNER	COAX	RCLD	S	S	TROOPS	VAL	N	GPI	COAX	<900
	211	GUNNER	COAX	RCLD	S	S	TROOPS	VAL	N	TEL	COAX	<900
	212	GUNNER	COAX	RCLD	S	S	TROOPS	VAL	N	INF	COAX	<900
	205	GUNNER	COAX	RCLD	S	M	TSV/CREW	VAL	N	GPD	COAX	<900
	206	GUNNER	COAX	RCLD	S	M	TSV/CREW	VAL	N	GPI	COAX	<900
	207	GUNNER	COAX	RCLD	S	M	TSV/CREW	VAL	N	TEL	COAX	<900
	200	GUNNER	COAX	RC	S	S	TSV	NVIS	D/N	AUX	COAX	<900
Cluster No. 13	142	TC	COAX	NP	S	S	TSV	VIS	D/N	RFD	COAX	<900
	143	TC	COAX	NP	M	S	TSV	VIS	D/N	RFD	COAX	<900
	144	TC	COAX	NP	S	S	TSV	VAL	N	RFD	COAX	<900
	145	TC	COAX	NP	S	S	TSV	VAL	N	RFI	COAX	<900
	146	TC	COAX	NP	M	S	TSV	VAL	N	RFD	COAX	<900
	147	TC	COAX	NP	M	S	TSV	VAL	N	RFI	COAX	<900
	148	TC	COAX	NP	S	M	TSV	VIS	D/N	RFD	COAX	<900
	149	TC	COAX	NP	M	M	TSV/CREW	VIS	D/N	RFD	COAX	<900
	150	TC	COAX	NP	S	M	TSV/CREW	VAL	N	RFD	COAX	<900
	151	TC	COAX	NP	S	M	TSV/CREW	VAL	N	RFI	COAX	<900
	152	TC	COAX	NP	M	M	TSV/CREW	VAL	N	RFD	COAX	<900
	153	TC	COAX	NP	M	M	TSV/CREW	VAL	N	RFI	COAX	<900
	154	TC	COAX	NP	S	S	TROOPS	VIS	D/N	RFD	COAX	<900
	155	TC	COAX	NP	M	S	TROOPS	VIS	D/N	RFD	COAX	<900
	156	TC	COAX	NP	S	S	TROOPS	VAL	N	RFD	COAX	<900
	157	TC	COAX	NP	S	S	TROOPS	VAL	N	RFI	COAX	<900
	158	TC	COAX	NP	M	S	TROOPS	VAL	N	RFD	COAX	<900
	159	TC	COAX	NP	M	S	TROOPS	VAL	N	RFI	COAX	<900
	281	TC	COAX	NP	MH	S	TSV	VIS	D/N	RFD	COAX	<900
	286	TC	COAX	NP	MH	S	TSV	VAL	N	RFI	COAX	<900
	287	TC	COAX	NP	MH	S	TSV	VAL	N	RFD	COAX	<900
	291	TC	COAX	NP	MH	M	TSV/CREW	VIS	D/N	RFD	COAX	<900
	296	TC	COAX	NP	MH	M	TSV/CREW	VAL	N	RFI	COAX	<900
	297	TC	COAX	NP	MH	M	TSV/CREW	VAL	N	RFD	COAX	<900

(113)

	OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
Cluster No. 14	213	TC	COAX	RCLD	S	S	TSV	VAL	N	RFD	COAX	<900
	214	TC	COAX	RCLD	S	S	TSV	VAL	N	RFI	COAX	<900
	215	TC	COAX	RCLD	S	M	TSV/CREW	VAL	N	RFD	COAX	<900
	216	TC	COAX	RCLD	S	M	TSV/CREW	VAL	N	RFI	COAX	<900
	217	TC	COAX	RCLD	S	S	TROOPS	VAL	N	RFD	COAX	<900
	218	TC	COAX	RCLD	S	S	TROOPS	VAL	N	RFI	COAX	<900
Cluster No. 15	303	TC	CAL.50	NP	M	S	LAV	VAL	N	TPI	CAL.50	<1000
	321	TC	CAL.50	NP	M	S	AIR	VAL	N	TPI	CAL.50	<1000
	300	TC	CAL.50	NP	M	S	LAV/CREW	VAL	N	TPI	CAL.50	<1000
	236	TC	CAL.50	NP	M	S	TROOPS	VAL	N	TPI	CAL.50	<1000
	317	TC	CAL.50	NP	S	S	AIR	VAL	N	TPI	CAL.50	<1000
	315	TC	CAL.50	NP	M	M	AIR	VAL	N	TPI	CAL.50	<1000
	222	TC	CAL.50	NP	S	S	LAV/CREW	VAL	N	TPI	CAL.50	<1000
	234	TC	CAL.50	NP	S	S	TROOPS	VAL	N	TPI	CAL.50	<1000
	224	TC	CAL.50	NP	MH	S	LAV/CREW	VAL	N	TPI	CAL.50	<1000
	306	TC	CAL.50	NP	MH	S	TROOPS	VAL	N	TPI	CAL.50	<1000
	319	TC	CAL.50	NP	MH	S	AIR	VAL	N	TPI	CAL.50	<1000
	313	TC	CAL.50	NP	MH	M	AIR	VAL	N	TPI	CAL.50	<1000
	228	TC	CAL.50	NP	S	M	LAV	VAL	N	TPI	CAL.50	<1000
	230	TC	CAL.50	NP	MH	M	LAV	VAL	N	TPI	CAL.50	<1000
	311	TC	CAL.50	NP	S	M	AIR	VAL	N	TPI	CAL.50	<1000

Cluster No. 16

OBJECTIVE NUMBER	CREW MEMBER	WEAPON	FIRING MODE	FIRING VEHICLE MOTION	TARGET MOTION	TARGET TYPE	TARGET VISIBILITY	DAY OR NIGHT	FIRE CONTROL INSTRUMENT	AMMUNITION	TARGET RANGE
219	TC	CAL.50	NP	S	S	LAV/CREW	VIS	D/N	TPD	CAL.50	<1600
221	TC	CAL.50	NP	S	S	LAV/CREW	VAL	N	TPD	CAL.50	<1600
231	TC	CAL.50	NP	S	S	TROOPS	VIS	D/N	TPD	CAL.50	<1600
233	TC	CAL.50	NP	S	S	TROOPS	VAL	N	TPD	CAL.50	<1600
225	TC	CAL.50	NP	S	M	LAV	VIS	D/N	TPD	CAL.50	<1600
227	TC	CAL.50	NP	S	M	LAV	VAL	N	TPD	CAL.50	<1600
237	TC	CAL.50	NP	S	M	AIR	VIS	D/N	TPD	CAL.50	<2300
310	TC	CAL.50	NP	S	M	AIR	VAL	N	TPD	CAL.50	<2300
316	TC	CAL.50	NP	S	S	AIR	VAL	N	TPD	CAL.50	<2300
220	TC	CAL.50	NP	MH	S	LAV/CREW	VIS	D/N	TPD	CAL.50	<1600
223	TC	CAL.50	NP	MH	S	LAV/CREW	VAL	N	TPD	CAL.50	<1600
304	TC	CAL.50	NP	MH	S	TROOPS	VIS	D/N	TPD	CAL.50	<1600
305	TC	CAL.50	NP	MH	S	TROOPS	VAL	N	TPD	CAL.50	<1600
226	TC	CAL.50	NP	MH	M	LAV	VIS	D/N	TPD	CAL.50	<1600
229	TC	CAL.50	NP	MH	M	LAV	VAL	N	TPD	CAL.50	<1600
232	TC	CAL.50	NP	M	S	TROOPS	VIS	D/N	TPD	CAL.50	<1600
235	TC	CAL.50	NP	M	S	TROOPS	VAL	N	TPD	CAL.50	<1600
309	TC	CAL.50	NP	M	S	AIR	VIS	D/N	TPD	CAL.50	<2300
320	TC	CAL.50	NP	M	S	AIR	VAL	N	TPD	CAL.50	<2300
314	TC	CAL.50	NP	M	M	AIR	VAL	N	TPD	CAL.50	<2300
322	TC	CAL.50	NP	M	M	AIR	VIS	D/N	TPD	CAL.50	<2300
298	TC	CAL.50	NP	M	S	LAV/CREW	VIS	D/N	TPD	CAL.50	<1600
302	TC	CAL.50	NP	M	S	LAV	VAL	N	TPD	CAL.50	<1600
301	TC	CAL.50	NP	M	S	LAV	VIS	D/N	TPD	CAL.50	<1600
299	TC	CAL.50	NP	M	S	LAV/CREW	VAL	N	TPD	CAL.50	<1600
238	TC	CAL.50	NP	MH	S	AIR	VIS	D/N	TPD	CAL.50	<2300
312	TC	CAL.50	NP	MH	M	AIR	VAL	N	TPD	CAL.50	<2300
307	TC	CAL.50	NP	MH	S	AIR	VIS	D/N	TPD	CAL.50	<2300
318	TC	CAL.50	NP	MH	S	AIR	VAL	N	TPD	CAL.50	<2300

APPENDIX B ADDITIONAL ASSUMPTIONS REGARDING WEAPONS, RANGING, FIRE CONTROL INSTRUMENTS, FIRING VEHICLE AND TARGET MOTION, AND TARGET RANGE FOR THE GENERATION OF THE TANK GUNNERY DOMAIN

ASSUMPTIONS

Ranging

1. When firing using the range-card-lay-to-direct-fire mode, ranging is required, regardless of range, once the target is illuminated with either white light or illuminating shell.

2. The Tank Commander will not range on main gun targets at night using the metascope attached to the rangefinder.

3. Ranging and tracking are right-handed operations for the Tank Commander and cannot be performed at the same time. When firing from a moving vehicle or at a moving target, the Tank Commander is therefore not ranging.

Weapon

1. When firing the coaxial machinegun in the battlesight mode, the Tank Commander will announce "Coax, Battlesight" in the ammunition element of the fire command.

2. When firing the coaxial machinegun, the Tank Commander will estimate range.

Fire Control Instrument

1. Although aim-off is only applied at ranges beyond 1600 meters, aim-off is considered a relevant task element to be included in those job objectives where the range given in the job objectives is either 500 to 3200 meters or 1100 to 3200 meters.

2. When the Tank Commander is firing the main gun because the gunner cannot identify the target, the gunner is using the sight that is appropriate for the target engagement (i.e., periscope for SABOT or HEAT, telescope for HEP or BEEHIVE).

PRECEDING PAGE BLANK

Firing Vehicle Motion

1. Targets to be engaged using battlesight are engaged while the firing vehicle is stationary, on the move, or from a brief halt.

2. Targets to be engaged using precision gunnery are engaged while the firing vehicle is stationary, or has come to a brief halt.

3. Targets to be engaged using range card data are engaged from the static position at which the range card was prepared.

4. When the Tank Commander is firing at a moving target using range card lay to direct fire, he fires using the range indexed into the rangefinder.

Target Motion

1. Troops are considered to be stationary area targets for machinegun engagements, while bunkers and crew-served weapons and light-armored vehicles are considered stationary point targets.

Range

1. Because of difficulties in ranging, main gun engagements occur at ranges $\geq$ 500 meters.

2. The coax machinegun ranges are $\leq$ 900 meters.

3. Infrared engagements are conducted at a range of $\leq$ 1000 meters between the target and the firing vehicle (and the target and illuminating vehicle).

4. The ranges for battlesight HEAT engagements are $\leq$ 1100 meters.

5. Beehive precision engagements are fired at ranges $\leq$ 1200 meters.

6. The ranges for battlesight SABOT engagements are $\leq$ 1600 meters.

7. The ranges for caliber .50 engagements with the TC periscope and non-aircraft targets are restricted to $\leq$ 1600 meters.

8. HEP engagements of moving targets are limited to ranges ≤ 1600 meters.

9. The ranges for caliber .50 engagements with the TC periscope and aircraft targets are restricted to ≤ 2300 meters.

10. Gunner periscope precision and RCLDF engagements can be fired at ranges ≤ 4400 meters.

11. Telescope precision engagements can be fired at ranges ≤ 3200 meters.

12. Telescope RCDLF SABOT/HEAT engagements can be fired at ranges ≤ 4400 meters.

13. Rangefinder engagements can be fired at ranges ≤ 4400 meters.

14. Night engagements are fired with white-light or flares at ranges ≤ 1600 meters and flares at ranges ≤ 4400 meters.

APPENDIX C BEHAVIORAL ELEMENTS (STEPS) USED TO DESCRIBE TANK GUNNERY JOB OBJECTIVES

BEHAVIORAL ELEMENTS - DESCRIPTIVE LABELS

Tank Commander Elements

1	TC ANNOUNCES "GUNNER"
*2	TC ANNOUNCES "GUNNER, DIRECT FIRE"
3	TC ANNOUNCES "CALIBER FIFTY"
4	TC PLACES CALIBER FIFTY SAFETY IN FIRE POSITION
5	TC PLACES CALIBER FIFTY SELECTOR IN HIGH RATE OF FIRE POSITION
6	TC ANNOUNCES "BATTLESIGHT"
7	TC ANNOUNCES "SABOT" OR "HEAT"
8	TC ANNOUNCES "HEP"
9	TC ANNOUNCES "BEEHIVE TIME"
10	TC ANNOUNCES "AREA FIRE" OR "ONE, TWO....ROUNDS, HEP"
11	TC ANNOUNCES "INDEX HEP, FIRE SABOT" OR "INDEX HEP, FIRE HEAT"
12	TC ANNOUNCES "COAX"
*13	TC ANNOUNCES "WHITELIGHT"
14	TC ANNOUNCES "REDLIGHT"
15	TC LAYS GUN FOR DIRECTION
16	TC ANNOUNCES "INDEX HEP, FIRE BEEHIVE" OR "BEEHIVE, TIME"
17	TC INDEXES RANGE INTO RANGEFINDER
18	TC ANNOUNCES TARGET DESCRIPTION
19	TC ANNOUNCES "MOVING"
20	TC ANNOUNCES "DRIVER, STOP"
21	TC RANGES
22	TC ANNOUNCES RANGE
23	TC ANNOUNCES DEFLECTION
24	TC VERIFYS DEFLECTION READBACK
25	TC ANNOUNCES QUADRANT ELEVATION
26	TC VERIFYS QUADRANT ELEVATION READBACK
27	TC LAYS CROSSHAIR AT CENTER OF BASE OF TARGET
28	TC LAYS CROSSHAIR AT CENTER OF TARGET VULNERABILITY
29	TC APPLIES AIM-OFF (RANGE > 1600 METERS)
30	TC APPLIES LEAD IN DIRECTION OF OWN GUN TRAVERSE
31	TC APPLIES LEAD IN DIRECTION OF TARGET APPARENT MOTION
32	TC LAYS CROSSHAIR LEADLINE AT CENTER OF BASE OF TARGET
33	TC LAYS RANGELINE AT CENTER OF TARGET VULNERABILITY
34	TC LAYS RANGELINE LEADLINE AT CENTER OF TARGET VULNERABILITY
35	TC ANNOUNCES "FIRE" OR "AT MY COMMAND. ...FIRE"
36	TC ANNOUNCES "FROM MY POSITION"
37	TC MAKES FINAL PRECISE LAY
38	TC TIMES SHOT
39	TC ANNOUNCES "ON THE WAY"
40	TC FIRES MAIN GUN

* Not currently used.

PRECEDING PAGE BLANK

Tank Commander Elements (Cont'd.)

41 TC FIRES COAX
42 TC FIRES CALIBER FIFTY
43 TC LAYS CROSSHAIR LEADLINE AT CENTER OF TARGET VULNERABILITY
44 TC ADJUSTS CALIBER FIFTY MACHINEGUN BURST FOR POINT TARGETS
45 TC ADJUSTS CALIBER FIFTY MACHINEGUN BURST FOR AREA TARGETS
46 TC ADJUSTS COAXIAL MACHINEGUN BURST FOR POINT TARGETS
47 TC ADJUSTS COAXIAL MACHINEGUN BURST FOR AREA TARGETS
48 TC FIRES CONTINUOUS BURST AT AIRCRAFT TARGETS
49 TC TURNS RANGEFINDER RETICLE SIGHT ON
50 TC ESTIMATES RANGE

51 TC ADDS 5 MILS ELEVATION (RANGE > 400 METERS)
52 TC LAYS COAX FOR DIRECTION ON EDGE OF TARGET
53 TC TURNS INFRARED SIGHT ON
54 TC LAYS M85 FOR DIRECTION ON EDGE OF TARGET

Gunner Elements

55 GUNNER TURNS ON TURRET POWER
56 GUNNER TURNS ON MAIN GUN SWITCH
57 GUNNER TURNS ON COAX SWITCH
58 GUNNER INDEXES HEP
59 GUNNER INDEXES APERS
60 GUNNER SELECTS SABOT OR HEAT RETICLE
61 GUNNER SELECTS HEP RETICLE
62 GUNNER ANNOUNCES "IDENTIFIED"
63 GUNNER ANNOUNCES "CANNOT IDENTIFY"
64 GUNNER TRAVERSES TO ANNOUNCED DEFLECTION

65 GUNNER READS BACK DEFLECTION
66 GUNNER INDEXES ANNOUNCED QUADRANT ELEVATION
67 GUNNER READS BACK QUADRANT ELEVATION
68 GUNNER LEVELS BUBBLE
69 GUNNER REINDEXES SABOT OR HEAT
70 GUNNER REINDEXES BEEHIVE
71 GUNNER LAYS CROSSHAIR AT CENTER OF BASE OF TARGET
72 GUNNER LAYS RANGELINE AT CENTER OF BASE OF TARGET
73 GUNNER LAYS CROSSHAIR AT CENTER OF TARGET VULNERABILITY
74 GUNNER LAYS RANGELINE AT CENTER OF TARGET VULNERABILITY

Gunner Elements (Cont'd.)

75 GUNNER APPLIES AIM-OFF (RANGE > 1600 METERS)
76 GUNNER LAYS CIRCLE RETICLE AT CENTER OF TARGET
77 GUNNER APPLIES LEAD IN DIRECTION OF OWN GUN TRAVERSE
78 GUNNER APPLIES LEAD IN DIRECTION OF TARGET APPARENT MOTION
79 GUNNER LAYS CROSSHAIR LEADLINE AT CENTER OF BASE OF TARGET
80 GUNNER LAYS RANGELINE LEADLINE AT CENTER OF BASE OF TARGET
81 GUNNER LAYS CROSSHAIR LEADLINE AT CENTER OF TARGET VULNERABILITY
82 GUNNER LAYS RANGELINE LEADLINE AT CENTER OF TARGET VULNERABILITY
83 GUNNER LAYS CIRCLE RETICLE AT INTERPOLATED LEAD POINT FOR TARGET
84 GUNNER MAKES FINAL PRECISE LAY

85 GUNNER TIMES SHOT
86 GUNNER ANNOUNCES "ON THE WAY"
87 GUNNER FIRES MAIN GUN
88 GUNNER FIRES COAX
89 GUNNER ADJUSTS COAXIAL MACHINEGUN BURST FOR POINT TARGETS
90 GUNNER ADJUSTS COAXIAL MACHINEGUN BURST FOR AREA TARGETS
91 GUNNER ANNOUNCES "SABOT INDEXED" OR "HEAT INDEXED" OR "BEEHIVE INDEXED"
92 GUNNER TURNS INFRARED SIGHT ON
93 GUNNER ADDS 5 MILS ELEVATION (RANGE > 400 METERS)
94 GUNNER LAYS COAX FOR DIRECTION ON EDGE OF TARGET

Driver Elements

95 DRIVER MAINTAINS STEADY RATE OF SPEED
96 DRIVER MANEUVERS TANK FOR FIRING
97 DRIVER ANNOUNCES ADVERSE TERRAIN CONDITIONS
98 DRIVER MOVES TO HULL DOWN FIRING POSITION
99 DRIVER BRINGS TANK TO A SMOOTH GRADUAL HALT
100 DRIVER BRAKES

Loader Elements

101 LOADER UNLOCKS AMMO READY RACK
102 LOADER SELECTS SABOT OR HEAT
103 LOADER SELECTS HEP
104 LOADER SELECTS BEEHIVE
105 LOADER UNLOADS MAIN GUN
106 LOADER LOADS MAIN GUN
107 LOADER PLACES MAIN GUN SAFETY IN FIRE POSITION

Loader Elements (Cont'd.)

108 LOADER PLACES COAX SAFETY IN FIRE POSITION
109 LOADER ANNOUNCES "UP"
110 LOADER ANNOUNCES "HEP, UP"
111 LOADER ANNOUNCES "BEEHIVE, UP"
112 LOADER SETS BEEHIVE FUZE SETTING TO ANNOUNCED RANGE
113 LOADER STOWS ROUND
114 LOADER MOVES TO POSITION FOR OBSERVATION

APPENDIX D METHOD OF CLUSTER ANALYZING THE JOB-OBJECTIVE/BEHAVIORAL-ELEMENT DATA MATRIX

Unlike many more familiar analytical techniques which are based on a general statistical model, cluster analysis is defined algorithmically. Algorithms are employed for two phases of the analysis. First, an algorithm must be selected for measuring the distance between rows of the matrix (in the present case, job objectives) as a function of the columns (behavioral elements). Second, an algorithm must be chosen for joining similar objectives into families, and partitioning dissimilar ones.

The selection of a distance algorithm is often difficult in cluster analysis, but was trivial in the present instance. The objective-element matrix contained only "1"'s and "0"'s, a case in which it can be shown, with few exceptions, that all common distance measures reduce to a single measure. Thus, the decision was made to use the inverse of the simple matching coefficient (SMC), one of the simplest variants of this common distance measure.

The SMC measured behavioral similarity in the present analyses in terms of the proportion of elements which were identical between each pair of objectives. Thus, for two objectives which had exactly the same values (1 or 0) on 20 out of the 114 elements, the inter-objective similarity was 20/114 or .175, and the inter-objective distance was 1.0-.175 = .825.

The selection of a joining algorithm was more difficult. Commonly used algorithms begin by joining the two closest (or identical) objectives into a single cluster. The next closest objective is then examined; if close enough to the first two (as defined by an arbitrary distance threshold), it is added to the cluster (a process called amalgamation); if not, it forms the seed of a new cluster. This process continues until all objectives have been examined once, and have been partitioned into many small clusters. The algorithm then passes through the data again, with a larger distance threshold, this time clustering the many small clusters into larger ones. The process goes through a number of cycles until all objectives have been amalgamated into one large cluster. Tracing the amalgamation process provides information on just how large the distances were between each pair of clusters before they were joined.

During this examination of the cluster solution or "tree", the analyst typically chooses a critical inter-cluster distance (or more than one for different regions of the data matrix) at which to anchor his interpretation of the solution; the set of clusters thus defined would be used to describe a stratification of the original matrix. When such clusters are inspected, one will find that the pattern of elements from objectives within a given cluster will always be more similar or homogeneous than the pattern of elements from objectives chosen across clusters. More will be said about interpretation later in this appendix.

The type of common algorithm described above is referred to as a direct joining algorithm, and has often been employed by numerical taxonomists. Recently, however, a series of difficulties with such algorithms have been pointed out, especially when employed with the kind of data of concern in this project (Hartigan, 1972). As one example, the SMC gives equal weight to each behavioral element, and the direct joining algorithm depends on and demands this equivalence in order to accomplish the amalgamation of objectives. However, in the present case such equal weighting implies that each behavioral element is equally important in determining performance of a job objective, and that each contributes equally to the generalizability of performance. Consider as an example the following two sets of elements:

Set 1	Set 2
TC announces "Gunner".	TC lays gun for direction.
TC announces "SABOT".	TC ranges.
TC announces target description.	TC indexes range into rangefinder.

When the three elements on the right are compared to the three on the left in terms of relative complexity, it is clear they should not be considered equivalent in their contribution to job objective performance. The fact is that a fire command, composed of the kinds of elements listed on the left, is always given for every engagement; and the skills involved in these individual pieces of the command are just not of the same magnitude as the skills involved in the perceptual and motor operations listed on the right. In cases like these where differential weighting of the elements may be of potential value, Hartigan recommends retaining the SMC data, in lieu of trying to subjectively weight the elements ab initio, and employing other joining algorithms which differentially weight elements during the amalgamation process.

Another example of the difficulties surrounding use of the direct joining algorithm is the required assumption of orthogonality among the columns (behavioral elements) of the matrix. This does not rule out empirical correlations among elements across objectives, but rather logical correlations. The ability to meet this assumption seemed questionable in the present case, since certain crew procedures reflected in several of the job objectives involved some elements which had to occur together if they occurred at all. As an example, consider elements 95 and 96 (see Appendix C). During a moving engagement, the driver always "maintains a steady rate of speed" and "maneuvers the tank for firing." This relationship is not merely fortuitous, but is required logically by the task; thus the assumption of orthogonality is violated for at least these two elements.

In order to deal effectively with the kinds of problems described above, Hartigan (1972) has proposed the use of a new algorithm known as the two-way direct joining algorithm or the block clustering algorithm. While this algorithm is still based on the SMC distance measure, it clusters both rows and columns of the data matrix simultaneously, using the SMC distance between objectives to cluster objectives, and the distance between variables (defined as the proportion of objectives which include each pair of elements) to cluster behavioral elements. The key point is that as the analysis proceeds, the clustering of job objectives ceases to depend on distances derived from consideration of individual elements; rather, distance among objectives is expressed in terms of the proportion of clusters of elements which are shared among objectives. This solves the weighting problem discussed earlier, since elements which are highly intercorrelated form clusters which are then used to measure inter-objective distances. Thus, the three behavioral elements of a fire command, which were discussed above, might form a cluster, in which case inter-objective distances would be based on a composite element defined for the complete fire command. The composite element would receive a weight equal to single elements which, at least in that stage of the analysis, retained their individual identities. The equal weighting of a cluster of three elements and a single element effectively triples the importance of the single element.

A second advantage is that the assumption of orthogonality is not required by the two-way algorithm, since dependencies are identified as elements are clustered, and only clusters of elements (where the clusters are orthogonal) are used in defining the clusters of job objectives.

An added advantage of the two-way direct joining algorithm is that interpretation is relatively straightforward since each cluster of objectives is defined directly on clusters of behavioral characteristics. Not only is interpretation direct, but when the element clusters defining clusters of objectives are at definite variance with the analyst's judgment, ad hoc adjustments to the solution can be made which are more public and replicable than are adjustments based on other algorithms.

The actual program used to implement the two-way joining algorithm was BMDP3M from the Biomedical Computer Programs package (Dixon, 1975). This specific program speeds up the two-way joining process by choosing a single job objective to represent each cluster of objectives resulting from a given pass through the data; similarly, a single element is used to represent each new element cluster. Objectives and elements singled out in this fashion are referred to as leaders (or as pass leaders, since some leaders from preceding passes are dropped when the clusters they represent are joined to another cluster with an established leader).

This specific implementation of the two-way direct joining algorithm is not without some potential difficulties of its own. For example, the cluster partition is not always invariant under various reorderings of the input data (e.g., the first objective will always be a leader), and clusters among objectives which are input first will always be larger, since they get the first chance at cluster membership (Hartigan, 1975, pp 77-78).

These weaknesses are partially offset by adding a reordering of the data matrix as the first step in the analysis. For each element, the value which occurs more frequently (0 or 1) is found, and the elements are ordered according to these frequencies. For each job objective, the number of elements whose values are modal are counted (i.e., if an element takes on the value "1" for 50 objectives in the domain, and the value "0" for 216 objectives, its modal value is "0"), and objectives are ordered by this count. After this bivariate reordering is complete, frequently occurring data values tend to lie in the upper left corner of the matrix (Dixon, 1975).

In summary, both major types of joining algorithms had potential weaknesses. Since no computer program of the exact type desired was available, and since, to the knowledge of the project staff neither approach had ever been employed for the kind of requirements at hand in the present project, preliminary analyses were performed using both algorithms. (The direct joining solution was obtained using program BMDP2M, Dixon, 1975). Based on the small but important differences between the outcomes of these preliminary studies, the two-way algorithm was selected for the actual partitioning of the job objectives into homogeneous families.

A series of two-way cluster analyses was performed. Each analysis differed primarily in the size of the threshold distances used to determine when two objectives (or task elements) should be placed within the same cluster or kept apart. As mentioned before, with each pass through the data these threshold distances were automatically increased so that eventually all objectives were joined into a single large cluster. During these analyses it was discovered that manipulating the size of the threshold increments could affect the course of clustering, specifically the degree to which clusters in later passes would be formed of simple clusters from earlier passes. Since the hierarchical nature of cluster formation is critical to the interpretation process, it was necessary to manipulate the increment in threshold distances until the hierarchical trees had the simplest possible structure.

In the solution which was finally adopted, clusters which formed first differed by at most one element; these clusters were then joined when they differed by at most two elements; and so on, until a series of clusters had been formed which differed by seven or fewer behavioral elements. At this point larger steps were used (approximately six elements per step) until the process was complete. The resulting output was a tree diagram with highly similar job objectives nearest to one another at the base, and the clusters and thresholds at which amalgamation occurred appearing as branches of the tree. This tree diagram formed the basis for the definition of families of objectives.

It will be useful to consider in detail an example of how the resulting output was interpreted. Figure D-1 presents that portion of the overall solution which consists of all .50 caliber job objectives. On the left side of the diagram each engagement is described by its number and by its general characteristics, such as tank motion and target type. The clustering tree produced by the program appears on the right side of the diagram. Numbers found on the tree indicate the threshold distances (in terms of proportion of identical elements) at which each joining of clusters occurred. The italicized symbols (i.e., a,---,i) accompanying each amalgamation are for purposes of reference in the discussion which follows.

The threshold distance values in Figure D-1 correspond to the distances among the job objectives included in the branches of the tree. For example, consider the group of objectives included in the branch labelled a (job objective numbers 303, 321, 300, 236, 317, 315, 222, and 234). The value next to the a indicates that these eight job objectives form a cluster when the cluster distance threshold (i.e., the distance between each objective in the cluster) is .082. This value is derived by considering the proportion of elements which differ between each pair of objectives in the cluster; in this instance they have different (1 or 0) values on 9.4 (on the average) elements (9.4/114 = .082). Similarly, the cluster labelled a is made up of subclusters, one of which includes objectives 303, 321, 300, and 236; the average distance among these four objectives

Job Objective Number	Firing Vehicle Motion	Target Motion	Target Type	Target Visibility	Day or Night	Fire Control Instrument	Range
303	M	S	LAV	VAL	N	TPI	<1000
321	M	S	AIR	VAL	N	TPI	<1000
300	M	S	LAV/CREW	VAL	N	TPI	<1000
236	M	S	TROOPS	VAL	N	TPI	<1000
317	S	S	AIR	VAL	N	TPI	<1000
315	M	M	AIR	VAL	N	TPI	<1000
222	S	S	LAV/CREW	VAL	N	TPI	<1000
234	S	S	TROOPS	VAL	N	TPI	<1000
224	MH	S	LAV/CREW	VAL	N	TPI	<1000
306	MH	S	TROOPS	VAL	N	TPI	<1000
319	MH	S	AIR	VAL	N	TPI	<1000
313	MH	M	AIR	VAL	N	TPI	<1000
219	S	S	LAV/CREW	VIS	D/N	TPD	<1600
221	S	S	LAV/CREW	VAL	N	TPD	<1600
231	S	S	TROOPS	VIS	D/N	TPD	<1600
233	S	S	TROOPS	VAL	N	TPD	<1600
225	S	M	LAV	VIS	D/N	TPD	<1600
227	S	M	LAV	VAL	N	TPD	<1600
237	S	M	AIR	VIS	D/N	TPD	<2300
310	S	M	AIR	VAL	N	TPD	<2300
316	S	S	AIR	VAL	N	TPD	<2300
220	MH	S	LAV/CREW	VIS	D/N	TPD	<1600
223	MH	S	LAV/CREW	VAL	N	TPD	<1600
304	MH	S	TROOPS	VIS	D/N	TPD	<1600
305	MH	S	TROOPS	VAL	N	TPD	<1600
226	MH	M	LAV	VIS	D/N	TPD	<1600
229	MH	M	LAV	VAL	N	TPD	<1600
232	M	S	TROOPS	VIS	D/N	TPD	<1600
235	M	S	TROOPS	VAL	N	TPD	<1600
309	M	S	AIR	VIS	D/N	TPD	<2300
320	M	S	AIR	VAL	N	TPD	<2300
314	M	M	AIR	VAL	N	TPD	<2300
322	M	M	AIR	VIS	D/N	TPD	<2300
298	M	S	LAV/CREW	VIS	D/N	TPD	<1600
302	M	S	LAV	VAL	N	TPD	<1600
301	M	S	LAV	VIS	D/N	TPD	<1600
299	M	S	LAV/CREW	VAL	N	TPD	<1600
228	S	M	LAV	VAL	N	TPI	<1000
230	MH	M	LAV	VAL	N	TPI	<1000
311	S	M	AIR	VAL	N	TPI	<1000
238	MH	S	AIR	VIS	D/N	TPD	<2300
312	MH	M	AIR	VAL	N	TPD	<2300
307	MH	S	AIR	VIS	D/N	TPD	<2300
318	MH	S	AIR	VAL	N	TPD	<2300

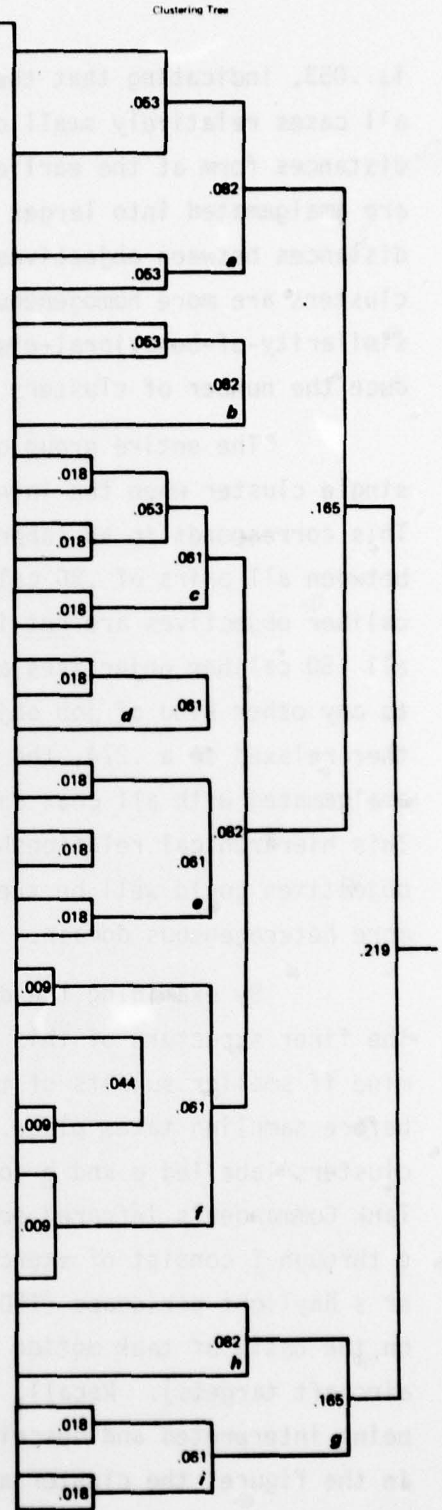

Figure D-1. Tree-clustering diagram for .50 caliber objectives.

is .053, indicating that they differ on an average of 6.0 elements. In all cases relatively small clusters with relatively small inter-objective distances form at the earliest branches of the tree (farthest left), and are amalgamated into larger clusters as the thresholds allow for greater distances between objectives. It is clear, therefore, that the early clusters are more homogeneous, and that one sacrifices homogeneity or similarity-of-behavioral-elements among objectives when attempting to reduce the number of clusters to a manageable size.

The entire group of .50 caliber objectives coalesces into a single cluster when the inter-objective threshold is relaxed to .219. This corresponds to an (approximate) average of 89 elements in common between all pairs of .50 caliber objectives. Further, since non-.50 caliber objectives are _not_ included in this cluster, one may argue that all .50 caliber objectives are more similar to one another than they are to any other kind of job objective. When the threshold distance is further relaxed to a .274, the super cluster of .50 caliber exercises is amalgamated with all coax and most main gun battlesight job objectives. This hierarchical relationship indicates that the group of .50 caliber objectives could well be treated as a homogeneous subsample taken from a more heterogeneous domain.

By examining the details of Figure D-1 it is possible to uncover the finer structure of this part of the cluster solution, and to determine if smaller subsets of the .50 caliber tasks should be broken out before sampling takes place. For example, it will be noted that the two clusters labelled _a_ and _b_ consist of exercises which are fired with the Tank Commander's infrared periscope (TPI), while the clusters labelled _c_ through _f_ consist of exercises which are all fired with the Tank Commander's daylight periscope (TPD). These clusters may be further subdivided on the basis of tank motion and target type (primarily aircraft vs. non-aircraft targets). Recall, however, that while the cluster structure is being interpreted and described in terms of the objective descriptors listed in the figure, the cluster analysis was actually performed on the much more detailed behavioral elements or steps required by each objective.

Thus, the tree-clustering diagram in Figure D-1 indicates that when one uses the infrared periscope instead of the daylight periscope, the pattern of behaviors involved changes; and this change is in some sense more significant than the change which occurs when one faces a moving target instead of a stationary target. In other words, it is assumed that TPI engagements are relatively poorer predictors of TPD engagements than are moving target engagements of stationary target engagements. Thus, there is a potential gain in predictive accuracy if the .50 caliber cluster is broken down into two components defined by TPI and TPD objectives. There is somewhat less benefit in any further breakdown by motion or target type.

The decision of when to stop subdividing clusters depends ultimately on practical constraints (e.g., the test should include at least one item from each cluster, so test length becomes important). The smaller clusters are certainly more homogeneous, as indicated by the smaller distance thresholds in the figure, but for a given cost a point of diminishing returns can usually be identified.

Referring once again to Figure D-1, one will notice that the seven job objectives at the bottom of the figure in cluster _g_ do not follow the pattern described in the preceding paragraphs (i.e., division between TPI and TPD objectives). If this last cluster is an artifact of the specific clustering solution used, it must be corrected for _ad hoc_ before interpretation of the cluster structure is finalized. Recall that the two-way joining algorithm used in these analyses surmounts the element weight problem by clustering objectives and elements simultaneously. As already indicated, however, this particular procedure is somewhat sensitive to the order in which the data are entered, and to particular combinations of rarely occurring elements. The program used in this study rotates the data matrix using a frequency criterion before the clustering process begins, and this exacerbates the problem of sensitivity to rarely occurring elements. Thus, one side effect of the rotation process is an occasional division of a homogeneous cluster into two pieces, where the second piece (in this case cluster _g_) is characterized by objectives which involve combinations of infrequent elements. In this instance, the elements associated with the combination of relatively infrequent aircraft targets, moving targets, and firing vehicles moving-to-a-halt produce the "outlier" cluster at the bottom of the figure. Notice that while this cluster is broken off from the rest of the .50 caliber engagements, it is itself divided as the rest are, into a TPI and a TPD cluster.

The decision of whether or not to integrate outlier components such as cluster g into other clusters requires the analyst to examine the underlying clusters of elements. In our example, the simple fact that there are fewer ways to engage aircraft targets than other kinds of targets may have produced the anomaly as an artifact of the algorithm. Since the two-way algorithm defines clusters directly on the elements involved, it is possible to examine unusual branches of this sort in the clustering tree to determine which anomalies are artifactual and to adjust them appropriately on an ad hoc basis. When the elements were examined in this case, it was found that the TPI engagements in cluster g were highly homogeneous with the other TPI engagements, and similarly the TPD engagements in g were very similar to those above, so the cluster structure for .50 caliber engagements was defined as having but two components--one containing all TPI objectives, and one containing all TPD objectives. This kind of ad hoc adjustment is often required in interpreting cluster analyses (Hartigan, 1975).

Finally, it should be noted that one final constraint was placed on interpretation of the cluster solution provided by the two-way joining algorithm. Although many of the primary conditions of engagement were differentiated quite clearly (e.g., by weapon, method of engagement, ammunition, etc.) the solution was occasionally insensitive to the crew member firing. For example, in the case of coax and main gun exercises there were clear distinctions in the solution as a function of ammunition and method of engagement, but occasionally gunner and tank commander exercises were amalgamated as the threshold increased. Specifically, gunner and tank commander coax exercises amalgamated before coax exercises in general amalgamated with the rest of the domain, and, similarly, gunner and tank commander battlesight and precision main gun engagements amalgamated across the person firing before they amalgamated across method of engagement. This is not surprising from the behavioral element point of view since, for given types of engagements, the activities performed by crew members are similar regardless of who actually fires the weapon. However, while the behavioral elements are similar, there is no reason that the performance of two distinct individuals should be similar. Thus, for testing purposes it was necessary to break the domain across the person firing (i.e., across gunner and tank commander) when the solution did not do so automatically. Had the gunnery domain been expressly analyzed for other purposes, such as generation of a new training program, the TC-gunner split possibly would not have received such emphasis.

APPENDIX E DERIVATION OF GENERALIZABILITY INDEX

Consider the hypotnetical family of objectives in Table E-1. This family consists of 10 objectives whose specific characteristics are defined in terms of eight behavioral elements. The largest number of elements involved in a single objective is six (out of the eight elements which are involved in the cluster at all). Three objectives (1, 4, and 10) have six out of the eight elements. Using the simple frequency approach one would select either objective 1, 4, or 10 to represent the family. Note, however, that some elements seem to be more characteristic of the family than others; for example, Element 8 is involved in all ten objectives in the cluster, while Element 3 is involved in only one. The cluster is defined in terms of those elements which most regularly occur across the cluster, and those elements may therefore be considered the most characteristic or representative. In other words, it is assumed that elements which have a high frequency across the cluster would account for a great deal of the performance variability which would result were all objectives in the cluster tested. Such elements therefore are most important in terms of generalizability of the unique performance aspects across the cluster. The row at the bottom of the table, labelled F_i, represents the number of times each element i occurs in the cluster, and represents an appropriate weight according to the argument above. Comparing the three job objectives 1, 4, and 10, it can be seen that while Objectives 1 and 4 both involve Element 1, Objective 10 does not, and, while Objective 10 involves Element 3, Objectives 1 and 4 do not. However, Element 1 occurs nine times in the cluster, while Element 3 occurs only once; therefore, it is relatively more important to represent Element 1 than Element 3 in that objective which is to be sampled as the test item.

Using this weight to generate an index of generalizability is quite straightforward. For each objective, simply add the weights (F_i) associated with each element contained in the objective. For example, this index, denoted ΣF_i, turns out to be 45 for Objective 1 (9+5+7+7+7+10), 44 for Objective 4, and 37 for Objective 10; the difference between 10 and the other two is largely because of the absence of Element 1 in Objective 10.

PRECEDING PAGE BLANK

Table E-1. DATA MATRIX FOR A HYPOTHETICAL CLUSTER OF JOB OBJECTIVES

Job Objective Number	Behavioral Elements 1	2	3	4	5	6	7	8
1	1	1	0	1	1	0	1	1
2	1	0	0	1	0	1	0	1
3	1	1	0	0	1	0	1	1
4	1	0	0	1	1	1	1	1
5	1	0	0	1	1	0	1	1
6	1	0	0	1	0	1	1	1
7	1	1	0	0	0	1	1	1
8	1	0	0	0	1	0	0	1
9	1	1	0	1	1	0	0	1
10	0	1	1	1	1	0	1	1
F_i	9	5	1	7	7	4	7	10
D_i	14	200	50	50	50	5	50	50
$\frac{F_i}{D_i}$	.64	.03	.02	.14	.14	.80	.14	.20
$F_i \bullet \frac{F_i}{D_i}$	5.76	.15	.02	.98	.98	3.2	.98	2

While use of F_i as a weight is attractive, another aspect of the data matrix must also be considered. Clearly, the more frequently an element occurs across the entire domain (D_i), the more frequently it will occur in any randomly selected cluster. To the extent that an element is frequent in a particular cluster and infrequent in the domain, it is even more important that such an element be represented in items selected for testing, since such elements describe behaviors which are unique to the particular kinds of engagements represented in the cluster. The ratio F_i/D_i reflects this uniqueness. At one extreme all occurrences of an element across the domain (D_i) might be found in a single cluster, yielding a ratio of 1.0. At the other extreme an element which is widely represented in the domain might not be found in a particular cluster, producing a ratio of 0.0. The domain frequencies and the F_i/D_i ratios are presented in Table E-1 for the hypothetical case. Using the ratio as a weight, an index of cluster-unique generalizability may be derived for each objective by calculating F_i/D_i for each element value (1 or 0) within a given objective and adding the products; this index is denoted $\Sigma F_i/D_i$.

Both of the weights described above had merit, so it was decided that they should be combined for actual sampling purposes. The combination was accomplished by multiplying the two weights together, i.e., $F_i \times (F_i/D_i)$, or F_i^2/D_i. This composite weight was then used to derive an index for each task by adding together the F_i^2/D_i values for each element involved in the task; this index was denoted $\Sigma F_i^2/D_i$. Table E-2 presents the three index values for each of the job objectives in the hypothetical cluster.

Table E-2. JOB OBJECTIVE INDEX VALUES FOR A HYPOTHETICAL CLUSTER

Job Objective Number	$\sum F_i$	$\sum F_i / D_i$	$\sum F_i^2 / D_i$
1	45	1.29	10.85
2	30	1.78	11.94
3	38	1.15	9.87
4	44	2.06	13.90
5	40	1.26	10.70
6	37	1.92	12.92
7	35	1.81	12.09
8	26	.98	8.74
9	38	1.15	9.87
10	37	.67	5.11

APPENDIX F CANDIDATE GUNNER MAIN GUN OBJECTIVES AND COMBINATIONS OF ENGAGEMENTS

Table F-1. CANDIDATE GUNNER MAIN GUN ENGAGEMENTS †

Number to draw	Objective Number	Crew Member	Weapon	Firing Mode	Firing Vehicle Motion	Target Motion	Target Type	Target Visibility	Day or Night	Fire Control Instrument	Ammo.	Target Range	Primary or Backup Method	Cluster Rank	Segment	Segment Rank
1	83	G	MG	RCLD	S	M	TNK/LAV	VAL	N	TEL	SH	< 4400	B	1	–	–
1	89*	G	MG	RCLD	S	S	TRPS	VAL	N	GPI	BEE	< 1000	P	1	–	–
1	244	G	MG	BS	MH	M	TNK/LAV	VIS	D/N	TEL	SH	< 1600	B	1	MH/	1
	252	G	MG	BS	MH	M	TNK/LAV	VAL	N	TEL	SH	< 1600	B	1	MH/	1
2	10	G	MG	BS	S	S	TNK/LAV	VAL	N	GPI	SH	< 1000	P	28	S/	1
	3	G	MG	BS	M	S	TNK/LAV	VIS	D/N	GPD	SH	< 1600	P	14	M/	2
	13	G	MG	BS	M	S	TNK/LAV	VAL	N	GPI	SH	< 1000	P	11	M/	1
	1	G	MG	BS	S	S	TNK/LAV	VIS	D/N	GPD	SH	< 1600	P	29	S/	4
2	40	G	MG	P	MH	M	TNK/LAV	VIS	D/N	TEL	SH	500-3200	B	1	/M	1
	43	G	MG	P	MH	S	TNK/LAV	VAL	N	GPD	SH	500-4400	P	12	/S	6
	48	G	MG	P	MH	M	TNK/LAV	VAL	N	TEL	SH	500-3200	B	1	/M	1
	35	G	MG	P	MH	S	TNK/LAV	VIS	D/N	GPD	SH	500-4400	P	9	/S	3
2	262	G	MG	P	MH	M	TSV	VIS	D/N	TEL	HEP	500-1600	P	1	/M	1
	63**	G	MG	P	MH	S	TRPS	VAL	N	GPD	BEE	500-4400	P	15	/S	4
	264	G	MG	P	MH	M	TSV	VAL	N	TEL	HEP	500-1600	P	1	/M	1
	59***	G	MG	P	MH	S	TRPS	VIS	D	GPD	BEE	500-4400	P	15	/S	4
9																
IMPLEMENTATION SUBSTITUTES																
*	274	G	MG	RCLD	S	M	TSV/CRW	VAL	N	GPI	HEP	< 1000	P	4	–	–
**	56	G	MG	P	MH	S	BKR/CRW	VAL	N	TEL	HEP	500-3200	P	3	/S	7
***	52	G	MG	P	MH	S	BKR/CRW	VIS	D	TEL	HEP	500-3200	P	3	/S	7

† See Table 2 for description of engagement conditions.

Table F-2

COMBINATIONS OF GUNNER MAIN GUN ENGAGEMENTS

Combination	Gunner Main Gun Objectives								
1.	83	89	244	10	3	40	43	262	63
2.	83	89	244	10	3	40	43	264	59
3.	83	89	244	10	3	48	35	262	63
4.	83	89	244	10	3	48	35	264	59
5.	83	89	244	13	1	40	43	262	63
6.	83	89	244	13	1	40	43	264	59
7.	83	89	244	13	1	48	35	262	63
8.	83	89	244	13	1	48	35	264	59
9.	83	89	252	1	3	40	43	262	63
10.	83	89	252	1	3	40	43	264	59
11.	83	89	252	1	3	48	35	262	63
12.	83	89	252	1	3	48	35	264	59
13.	83	89	244	10	3	40	35	262	63
14.	83	89	244	10	3	40	35	264	59
15.	83	89	244	10	3	40	43	262	59
16.	83	89	244	10	3	48	35	262	59
17.	83	89	244	13	1	40	35	262	63
18.	83	89	244	13	1	40	35	264	59
19.	83	89	244	13	1	40	43	262	59
20.	83	89	244	13	1	48	35	262	59
21.	83	89	252	1	3	40	35	262	63
22.	83	89	252	1	3	40	35	264	59
23.	83	89	252	1	3	40	43	262	59
24.	83	89	252	1	3	48	35	262	59