AD-A048 607 HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA F/G 5/9
CRITICALITY AND CLUSTER ANALYSES OF TASKS FOR THE M48A5, M60A1,--ETC(U)
NOV 77 J A BOLDOVICI, J H HARRIS, W C OSBORN DAHC19-76-C-0001
UNCLASSIFIED HUMRRO-FR-WD(KY)77-12 ARI-TR-77-A17 NL

1 OF 2
AD
A048607

ADA048607

(12)

AD

(19) ARI TECHNICAL REPORT

(18) ARI (15) TR-77-A17

(6) # CRITICALITY AND CLUSTER ANALYSES OF TASKS FOR THE M48A5, M60A1, AND M60A3 TANKS

(9) Final rept., 12 Apr 76 – 14 Feb 77,

by

(10) John A. Boldovici, James H. Harris, William C. Osborn and Charlotte L. Heinecke

(14) HumRRO-FR-WD(KY)77-12

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

(11) NOV [redacted] 77 (12) 179 p.

(15) Contract DAHC 19-76-C-0001

ARI Field Unit, Fort Knox, Kentucky

(16) 2Q762722A764

Prepared for

ari

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

405 260

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

J. E. UHLANER
Technical Director

W. C. MAUS
COL, GS
Commander

Research accomplished under
contract to the Department of the Army

Human Resources Research Organization
Western Division

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 Technical Report 77-A17	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CRITICALITY AND CLUSTER ANALYSES OF TASKS FOR THE M48A5, M60A1, and M60A3 TANKS		5. TYPE OF REPORT & PERIOD COVERED Final Report Task 1 of 2, 12 April 1976 - 14 Feb 1977
		6. PERFORMING ORG. REPORT NUMBER FR-WD(KY)-77-12
7. AUTHOR(s) John A. Boldovici, James H. Harris, William C. Osborn and Charlotte L. Heinecke		8. CONTRACT OR GRANT NUMBER(s) DAHC 19-76-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q762722A764
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Research Institute for the Behavioral and Social Sciences 5001 Eisehnower Avenue, Alexandira, VA 22333		12. REPORT DATE November 1977
		13. NUMBER OF PAGES 179
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		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 performed by HumRRO, Western Division, Fort Knox Office, P.O. Box 293, Fort Knox, Kentucky 40121; monitored technically by the Army Research Institute Field Unit at Fort Knox, under Donald F. Haggard, Chief.

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

Tanks, M48A5, M60A1, M60A3, Task Criticality, Paired Comparison Techniques, Inter-rater Reliability, Cluster Analysis, Learning Difficulty, Evaluation Difficulty.

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

The work reported here covers an analysis of armor crewman job tasks for the purpose of designing training for Reserve Components that use the M48A5 tank.

Task data were generated and organized for each tank crew position in a form that shows which tasks are common and unique to three tanks, M48A5, M60A1, and M60A3. Task criticality was estimated using a paired comparison rating technique in which raters selected hypothetical crewmen for a combat over

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

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Item 20.

misson, based on which tasks the crewmen could and could not perform. Reliability of the ratings averaged .68. Ways of improving the quality of task criticality studies were discussed.

Cluster analysis was used to group tasks by crew position according to similarities among descriptors by which the tasks were characterized. Eighty task clusters or "skills" were identified, 21 for the Driver, 19 for the Loader, 20 for the Gunner, and 20 for the Tank Commander.

Criticality learning difficulty and evaluation difficulty were estimated for each task cluster.

Results of the research indicated that: (1) The task analyses and the task criticality studies yielded results that will be useful for assigning training priorities; (2) the cluster analyses produced groups of tasks which appear reasonable, though the implications for training design remain to be demonstrated; and (3) results of the learning and evaluation difficulty studies were inconclusive.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

BRIEF

This report describes the conduct and results of the first task of a two-task project to design training for Armor and Cavalry National Guard units.

REQUIREMENT

The requirement to which Task 1 was addressed was to analyze tasks, estimate criticality, and perform related work in preparation for designing training for Reserve Components[1] that use the M48A5 tank. The objectives to be achieved during this preparatory work were to:.

1. Generate and organize task data for the M48A5, M60A1, M60A3, and XM-1 tanks.
2. Identify tasks that are common and unique to the M48A5, M60A1, and M60A3.
3. Use a paired-comparison technique to estimate the relative criticality of tasks for each of the three tanks.
4. Establish the reliability of the task criticality estimates.
5. Prepare plans for investigating the validity of the criticality estimates.
6. Use cluster analysis to group tasks into "skills," according to descriptors that have implications for training design.
7. Estimate the criticality, and the difficulty of learning and evaluating each of the task groups or "skills" identified as the result of item 6, above.

PROCEDURE AND RESULTS

Achieving the objectives listed above was described in four parts:

1. Generating and Organizing Task Data.
2. Task Criticality.
3. Cluster Analysis.
4. Skill Criticality, Learning Difficulty, and Evaluation Difficulty.

---

[1]"Reserve Components" as used in this report, refer to National Guard and U.S. Army Reserve units. With few exceptions, the only Reserve Components that are using or scheduled to use the M48A5 tank are Armor and Cavalry National Guard units.

Generating and Organizing Task Data

The project began with generating and organizing task data for the tank systems. Data sources included task data cards from the U.S. Army Armor School, research reports, operators' and equipment manuals, and task lists generated by the project staff. The task data were presented separately for each duty position in a form that shows which tasks are common and unique to the M48A5, M60A1, and M60A3.[1]

Task Criticality

Task criticality was estimated using a paired comparison study. Forty-eight AOAC (Armor Officers' Advanced Course) students selected hypothetical crewmen for a combat mission, based on which tasks the crewmen could and could not perform. The assumption here was that the officers' perceptions of task criticality would be reflected in their choices of crewmen to take into combat. The study yielded numerical indexes of criticality for each task.

The tasks receiving the highest criticality ratings were those that would be expected by one familiar with tank operations: the Tank Commander acquiring targets, the Tank Commander and Gunner firing the main gun, the Loader loading, and the Driver driving tactically.

The reliability of the paired comparison judgments was estimated by correlating the scale values of tasks common to the three tanks. Correlations, computed by duty position for each pair of tanks, ranged from .55 to .79, with an average of .68. All were statistically significant ($p < .05$).

Suggestions were offered as to how inter-rater reliability might be increased in future studies of task criticality with the paired comparison technique:

1. Increase the precision of defining the parameters on which judgments are to be made.
2. Provide opportunity for rater practice.

---

[1] Data for the XM-1 were submitted under separate cover. They were not used in later analyses because they were preliminary and subject to change.

3. Use complete, as opposed to partial, pairing designs.
4. Increase the number of observations per paired comparison.

A plan was presented for examining the construct validity of the criticality estimates. Issues associated with the content and predictive validity of criticality measurement also were discussed.

Cluster Analysis

Cluster analysis was used to group tasks according to similarities among descriptors by which the tasks were characterized. The exercise began with a search for a set of descriptors which could be used to characterize all armor tasks, and which might have implications for training design. Thirty-six descriptors were selected and used. Eleven of the 36 describe stimuli that initiate and maintain task performance; written materials and oral commands are examples. Six of the descriptors pertain to the tools, instruments, and controls that are used in task performance; variable setting controls, for example, and common hand tools. Eleven descriptors pertain to the mediating processes involved in task performance; using rules, for example, and recalling set procedures. The remaining eight descriptors describe overt responses; finger manipulation, for example, and reporting in writing.

The 36 descriptors were arrayed across the tops of data recording forms, with tasks and subtasks listed down the left margin. Two members of the project staff independently filled in the data tables, entering a "1" in the columns corresponding to descriptors that characterized each subtask, and leaving blank the descriptor columns that did not pertain to the subtask. The two sets of one-zero data thus generated served as the inputs for the inter-rater reliability studies that followed.

Inter-rater reliability was examined by computing phi (ϕ) coefficients for each of the four descriptor subsets (Stimuli; Tools, Instruments, and Controls; Mediating Process; and Overt Responses), and across subsets, both before and after rater practice. Doing so permitted examining not only inter-rater reliability, but also the effects of practice on inter-rater reliability.

Inter-rater reliability increased significantly with practice and discussion, irrespective of whether the tasks rated after practice were the same as or different from the tasks rated for practice. Overall inter-rater reliabilities for the tasks rated after practice were about .70.

After inter-rater reliability was examined, the two raters discussed their ratings, and produced a single, reconciled, task by task-descriptor matrix, which was the input for the cluster analyses.

The results of four cluster analyses, one for each duty position across the three tank systems, were presented. Eighty task clusters or "skills" were identified, 21 for the Driver, 19 for the Loader, 20 for the Gunner, and 20 for the Tank Commander. Examples of the skills for each duty position are:

1. Driver (M60A1, M48A5, M60A3), Perform Tank Operation Procedures: Performs fixed procedure multi-limb manipulation of various controls in response to oral commands.
2. Loader (M60A1, M48A5, M60A3), Perform Tactical Loading: Performs fixed procedure finger-hand-arm manipulation of various controls in response to oral commands by recalling information; reports by talking.
3. Gunner (M60A1, M48A5, M60A3), Perform Misfire Procedures: Performs fixed procedure finger-hand-arm manipulation of various controls in voluntary response to non-verbal sounds and body-feel while communicating orally.

4. Tank Commander (M60A1, M48A5, M60A3), Boresight and zero weapons: Performs continuous and fixed procedure finger-hand-arm manipulation of various controls and sometimes common hand tools in voluntary response to man-made environmental features, instrument read-outs and sometimes touch by recalling facts and classifying information; reports by talking.

The tasks comprising each of the 80 task clusters are listed by duty positions in Appendix B.

Skill Criticality, Learning Difficulty, and Evaluation Difficulty

Skill criticality, the mean of the criticality scores for the tasks comprising each of the 80 task clusters, was judged not particularly useful for training design.

Learning difficulty and evaluation difficulty for the domain of tank crew behavior associated with each task descriptor were rated by five members of the project staff. The estimates for each descriptor were averaged across raters. Difficulty estimates for each skill were then made by assigning the descriptor scores to the modal descriptor pattern for each skill.

The estimates of learning and evaluation difficulty were highly reliable (.76 and .88) in terms of the stability of the mean ratings obtained. The results were, however, judged inconclusive, because some seemed at odds with reality. The Driver's cluster, "Start Tank Engine," for example, received an extremely high difficulty rating. The apparent abberations may have been the result of deficiencies in the methods for computing difficulty, inappropriate naming of some clusters, or both.

Suggestions were made for examining the construct validity of learning and evaluation difficulty using designs similar to the one presented for criticality (Appendix F). Construct validity was tentatively examined in light of correlations between learning and evaluation difficulty ($r = .76$), and between each of the difficulty estimates and criticality ($r = .44$ in both cases).

USE OF FINDINGS

The results reported here are intended to be used during Task 2 to design training for Reserve Components that use the M48A5 tank. The task analyses and the task criticality studies yielded results that will be useful for assigning training priorities. The cluster analyses produced reasonable-appearing groups of tasks, though the implications for training design remain to be demonstrated. The results of the learning and evaluation difficulty studies were inconclusive, and will not be used.

PREFACE

This is the Final Report for Task 1 of a two-task project entitled "Tank Systems Skills and Training Structure." The report describes task-analytic and related work done in preparation for developing training outlines for Reserve Components that use the M48A5 tank.

The work reported in this volume was performed at the Fort Knox Office of the Human Resources Research Organization (HumRRO), under Contract No. DAHC-19-76-C-0001 with the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI).

John A. Boldovici is directing the project, which is staffed by Roy C. Campbell, J. Patrick Ford, James H. Harris, Charlotte L. Heinecke, Richard E. O'Brien, and William C. Osborn.

Paul W. Fingerman, Andrew M. Rose, and George R. Wheaton of the American Institutes for Research assisted substantially in interpreting the results of the cluster analysis under a subcontract with HumRRO.

Donald F. Haggard, the Contracting Officer's Technical Representative, provided administrative assistance, valuable criticism, and substantive suggestions for conceptualizing problems and solutions throughout the project.

The criticality study that was part of Task 1 could not have been conducted without the cooperation of many people. MAJ Douglas W. Smith, ARI Senior R&D Coordinator at Fort Knox, assisted in recruiting and scheduling subjects. Carolyn Harris assisted in designing the study. The officers who served as subjects were, as usual, gracious and cooperative.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd.)

LIST OF TABLES AND FIGURES

LIST OF TABLES AND FIGURES (Cont'd.)

CRITICALITY AND CLUSTER ANALYSES OF TASKS FOR THE M48A5, M60A1, AND M60A3 TANKS

The training needs of Reserve Components are changing. The M48A1 tank, which is the second most prevalent in the National Guard inventory, is being replaced by the M48A5. Personnel turbulence, always a problem in Reserve Components, promises to become even greater with the elimination of the draft, and as the result of expiration of the eight-year commitments of Guardsmen who entered service during the Vietnam build-up. In addition to problems associated with equipment and personnel turbulence, the costs of ammunition, real estate, range and hardware maintenance, targets, fuel, transportation, and replacement equipment continue to increase.

One effect of the trends noted above is that existing training for Armor and Cavalry Reserve Components is becoming increasingly inappropriate and obsolete. As old equipment is replaced with new, the training for operation and maintenance of the old equipment becomes inappropriate, and the need for new training becomes more compelling. As experienced Guardsmen are replaced with inexperienced personnel, training that focuses on higher level skills becomes insufficient, and training on basic skills becomes necessary. And as costs increase, training that depends on large quantities of ammunition, on frequent service practice firing, and on travel to and from training sites becomes less acceptable, and the need for training that can be delivered at armories becomes more obvious.

In the course of designing nearly any instructional program, several difficult problems must be solved. These include:

1. How to select tasks or objectives for inclusion in training.
2. How to group tasks for optimal efficiency of presentation in training.

A common method of selecting tasks for inclusion in training is to do so on the basis of task criticality; that is, to address only those tasks whose mastery is most critical to effective performance on the job. Measuring task criticality is, however, fraught with problems. Raters may not agree on which tasks are most critical (a reliability problem), and the ratings may be influenced by considerations other than criticality (a validity problem). If measuring criticality is unreliable, invalid, or both, then decisions about training content based on criticality measurement are bound to be in error.

Even if perfect reliability and validity were achieved in decisions about training content, the problem of bridging the gap between a task list and sets of tasks or objectives grouped for optimal presentation in training would remain. The issue of grouping tasks for training has been addressed indirectly in basic research on behavior classification and types of learning.[1] It has been addressed more directly in applied work on methods for training development,[2,3,4] usually as a prelude to selecting media, materials, and methods. Sorting tasks for presentation in training is necessarily a subjective matter, and little is known about the reliability of the results obtained. Adoption of the methods for sorting tasks has not been widespread, perhaps because users find implementation difficult. To the extent that methods for sorting tasks could be routinized, two benefits would seem to accrue: The methods might become easier to use, and the reliability of the results obtained might increase.

---

[1] See, for example, Gagné, R.M. The Conditions of Learning. New York, New York: Holt, Rinehart and Winston, 1965.

[2] Gropper, G.L., and Short, J.G., Handbook for Training Development, Pittsburgh, Pennsylvania: American Institutes for Research, 1969.

[3] Schumacher, S.P., and Glasgow, A.Z., Handbook for Designers of Instructional Systems, Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratories, 1973.

[4] US Army Transportation School. Interservice Procedures for Instructional Systems Development. Fort Eustis, Virginia: Author, 1975.

RATIONALE

Recognizing the dual need for new Reserve Component training and for addressing the training development issues outlined above, the US Army Research Institute for the Behavioral and Social Sciences (ARI) has undertaken research to:

1. Design training plans for operating and maintaining the M48A5 tank.
2. Explore new methods for establishing task criticality, and for grouping tasks for presentation in training.

This project is part of that research.

PURPOSE

The ultimate purpose of the project is to design training for Reserve and National Guard units that use M48A5 tanks. This report describes the work performed during Task 1, whose purposes were to:

1. Generate and organize task data for the M48A5, M60A1, M60A3, and XM-1 tanks.
2. Identify tasks that are common and unique to the M48A5, M60A1, and M60A3.
3. Use a paired-comparison technique to estimate the relative criticality of tasks for each of the three tanks.
4. Establish the reliability of the task criticality estimates.
5. Prepare plans for investigating the validity of the criticality estimates.
6. Use cluster analysis[1,2] to group tasks into "skills," according to descriptors that have implications for training design.
7. Estimate the criticality, and the difficulty of learning and evaluating each of the task groups or "skills" identified as the result of item 6, above.

---

[1]Hartigan, J.A. Direct clustering of a data matrix. Journal of the American Statistical Association, 67, 1972.

[2]Dixon, W.J., (Ed.). BMDP: Biomedical Computer Programs. Berkeley, California: University of California Press, 1975.

ORGANIZATION OF THE REPORT

How each of the objectives listed above was achieved is described in four major sections of the report:

1. **"Generating and Organizing Task Data" addresses the first and second objectives listed above.**
2. **"Task Criticality" addresses the third, fourth, and fifth objectives.**
3. **"Cluster Analysis" addresses the sixth objective.**
4. **"Skill Criticality, Learning Difficulty, and Evaluation Difficulty" addresses the seventh objective.**

GENERATING AND ORGANIZING TASK DATA

The project began with generating and organizing task data. The task lists would be used later in the project in a study of task criticality and in exploring the utility of cluster analysis as a method of grouping tasks for presentation in training.

Four tanks were addressed, in order to include systems used at present, and systems planned for use in the future:

1. The M60A1, which now predominates in the Active Army and National Guard.
2. The M60A3, an improved (retrofitted) version of the M60A1.
3. The M48A5, which is replacing the second most prevalent tank in the National Guard (the M48A1) and will thus become, with the M60A1, the "staple" for Reserve Components.
4. The XM-1, which eventually will become the US Army's main battle tank.

METHOD

Task lists for both XM-1 prototypes were written, using preliminary training outlines, equipment data, and manuals that were available at the time. The task lists have been presented elsewhere,[1] but were not used in later project work since the data were preliminary and subject to change.

Assembling the task data for the other three tanks began with a review of operations and maintenance tasks that had been rated critical or important in earlier studies by the US Army and its contractors. This preliminary task pool or data base was supplemented with tasks from a recent report on tank gunnery testing,[2] from operators' manuals and

---

[1]O'Brien, R.E., and Boldovici, J.A. Task Lists for Chrysler XM-1 Prototype (Project Memorandum No. 3). Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

[2]Boldovici, J.A., Wheaton, G.R., and Boycan, G.G. Selecting Items for a Tank Gunnery Test. Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

equipment data, and from additions based on local expertise. The sources for the task data are presented in Table 1, with summaries of the main differences between the M60A1 task list and the lists for the other two tanks. Additional details about generating and organizing the task data are presented in Appendix A.

RESULTS

Separate task lists for the M60A1, M48A5, and M60A3 were presented under separate cover.[1] A combined list, showing tasks that are common and unique to the three tanks, is presented in Appendix B. The cluster designations and criticality scores in Appendix B can be ignored now; they will be discussed later. Tasks in Appendix B that are common or unique to the three tank systems can be identified by either or both of two methods. The first two tasks in the Driver's list appear in Appendix B as:

		CRITICALITY		
TASK NO.	TASK	M60A1	M48A5	M60A3
AD105	Install the M27 periscope	5.355		4.402
A5111	Install the M27 periscope (spare)		4.348	

The first task (AD105) has entries in the criticality columns under M60A1 and M60A3, but not under M48A5. This indicates that the task is performed by M60A1 and the M60A3 Drivers, but not by M48A5 Drivers. The second task (A5111), has an entry in the criticality column under M48A5, but not under M60A1 or M60A3. This indicates that the task is performed by M48A5 Drivers, but not by M60A1 or M60A3 Drivers.

A less direct method of identifying tasks that are unique or common to the three tanks is by using the task code numbers (extreme left column of Appendix B). The codes are explained in Appendix C.

---

[1]Harris, J.H. Task Lists for M60A1, M60A1(AOS), M48A5, and M60A3 Tanks (Project Memorandum No. 1). Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

Table 1

DATA SOURCES FOR THE TASK LISTS, AND SUMMARY OF DIFFERENCES BETWEEN THE M60A1 TASK LIST AND THE TASK LISTS FOR THE OTHER TWO TANKS

	M60A1	M48A5	M60A3
DATA SOURCES	1. Task data cards (US Army Armor School, 1976). 2. Ford, Harris, & Rondiac (1974). 3. Boldovici, Wheaton, and Boycan (1976).	1. M60A1 task list. 2. M48A5 Operator's Manual (Hq., Dept. of Army, 1975). 3. Boldovici, Wheaton, and Boycan (1976).	1. M60A1 task list. 2. Boldovici, Wheaton, and Boycan (1976). 3. M60A1E3 Operator's Manual (Chrysler Corporation, 1974). 4. M60A1 Product Improvement (M60A1E3) (U.S. Army Material Command, 1975). 5. Phase II Product Improvements for M60A1/M60A1-P1 Tanks, (Anonymous, Undated). 6. Tank Thermal Sight, (Texas Instruments Incorporated, 1976).
MAIN DIFFERENCES FROM M60A1		Twenty-two tasks added, which staff judged important or critical, but which were not in 11E task list; e.g., "Check track tension," "Connect track," "Zero M2 machinegun."	Includes precision engagements from moving tanks. Includes tasks related to eight product improvements: laser rangefinder, electronic computer, light amplication sights, tank thermal sight, smoke grenade launcher, muzzle reference system, MAG-58 coaxial machinegun, Driver's viewer (VVS2).

TASK CRITICALITY

Training resource limitations demand that choices be made about what to include in training, and what to exclude. Agreement seems widespread that training programs should minimally include tasks that are critical to effective job performance (and cannot be performed by new trainees). In military training contexts, this reduces to including in training those tasks that are essential (critical) to effective performance in combat. Since combat cannot be realistically simulated, a measurement problem immediately arises; namely, how to measure criticality.

Prescriptive training development literature such as the Interservice Procedures for Instructional Systems Development[1] typically mentions task criticality as an important consideration in determining training content. The literature is, however, vague on the question of how to measure criticality, and silent on the measurement issues associated with criticality estimation.

Conventional training development methods deal with the problem of selecting tasks for inclusion in training in the following way: A job analysis is conducted, resulting in a task list or "inventory." Expert judgment is then used to rate the criticality of each task on some *n*-point scale ranging from "irrelevant to the job" to "highly critical to mission accomplishment." The tasks receiving the highest ratings are selected for inclusion in training, and those receiving low criticality ratings are excluded or deemphasized. Since the content of training frequently is determined on the basis of criticality ratings, a question naturally arises as to how much confidence can be placed in the ratings. One index of confidence is inter-rater reliability: to the extent that

[1]US Army Transportation School, op. cit., 1975.

several raters independently produce similar criticality ratings, confidence in the job-relevance of training content based on the ratings increases. The test-development axiom is directly analogous: reliability is necessary for validity. Applied to training content, the axiom becomes "reliability (of criticality ratings) is necessary for job-relevance (of training content)."

The reliability of criticality ratings that are used for determining training content seldom is reported.[1,2] In the few instances where reliability has been reported[3] rater agreement has been poor -- too low in fact for the ratings to be of practical use. An exception appears in a recent test-development project[4]: Two-hundred forty tank gunnery tasks were ranked in terms of criticality, which was determined by the use of a paired-comparison technique. The Tank Commanders serving as subjects were presented with many pairs of target/range combinations. (An example of a pair of target/range combinations is tank at 2000 to 2500 meters, and light-armored vehicle at 500 to 1000 meters.) The subjects were instructed to assume that they had encountered each pair of target/range combinations on the battlefield, and that they could not engage both targets simultaneously. They were then asked to indicate which one of the two target/range combinations that comprised each item they would engage first. A criticality score was computed by counting the number of times each combination was chosen as more threatening ("would be engaged first") and dividing by the number of times it could have been chosen.[5] Inter-rater reliability was in the high nineties.

---

[1]McCluskey, M.R., Jacobs, T.O., and Cleary, F.K. Systems Engineering of Training for Eight Combat Arms MOSs, Alexandria, Virginia: Human Resources Research Organization (HumRRO), 1975.

[2]McKnight, J.A. and Hundt, A.G. Driver Education Task Analysis: The Development of Instructional Objectives. Alexandria, Virginia: Human Resources Research Organization (HumRRO), 1972.

[3]Ammerman, H.L. and Pratzner, F.C. Occupational Survey on Auto Mechanics: Task Data from Workers and Supervisors Indicating Job Relevance and Training Criticalness. Columbus, Ohio: Ohio State University, 1975.

[4]Boldovici, J.A., Wheaton, G.R., and Boycan, G.G., op. cit., 1976.

[5]Guilford, J.P. Psychometric Methods. New York, New York: McGraw Hill, 1954.

Since the rated items varied only in target type and range, the judgments about target threat or criticality were easy to make. The high degree of rater agreement probably also reflected certain learning experiences that the subjects had in common: Tank Commanders receive formal instruction in assessing target threat. The high inter-rater reliability, therefore, may simply have indicated that all of the subjects had learned "the same things." One wonders then, whether similarly high inter-rater reliability could be achieved using the paired-comparison technique with a heterogeneous sample of tasks, where the dimensions for making the criticality judgments were less obvious than target type and range, and where the subjects had not received formal instruction in making judgments of the kind required for the ratings. The present study provided for answering the question.

PURPOSE

The purpose of the study was to use a paired comparison technique to estimate the relative criticality of armor tasks rated critical and important in earlier studies, and to establish the inter-rater reliability of the estimates produced in the present study.

METHOD

Respondents

Forty-eight captains, who were enrolled in the Armor Officers' Advanced Course (AOAC) at Fort Knox during the conduct of the study, served as respondents.

Questionnaires

Twelve forms of a paired comparison questionnaire were used. The units of comparison in each form were the tasks for one of four crew positions (Driver, Loader, Gunner, or Tank Commander) in one of three tanks (M60A1, M48A5, M60A3).

The design of each form of the questionnaire can be illustrated by describing how the form for the M60A1 Driver tasks was designed. Seventy M60A1 Driver tasks were identified during the task-description part of the project. The number of possible different pairs of 70 tasks is $70 \times 69/2 = 2415$. This would have been too many judgments for each respondent to make. A partial paired comparison design[1] was therefore used, in which each of the 70 tasks was paired with each of seven other tasks. The partial pairing yielded 245 unique pairs of tasks for the M60A1 Driver. The numbers of pairs of tasks for the other 11 forms of the questionnaire are shown in Table 2. Details of how the task pairs were formed are presented in Appendix D.

Procedure

The Captains who volunteered for participation in the study were instructed to be at a designated site at a particular time. Each of the first 12 to arrive was given a different form of the questionnaire. Each of the next 12 was given a different form, and so forth, until each of the 12 forms had been given to four respondents.

The respondents were instructed to assume that they were company commanders choosing crew members to take on a mission in which fire would be exchanged with the enemy. They were then asked to indicate which of two crew members they would choose, based on whether the crew member could do one or the other of a pair of tasks. An example of a pair of tasks for the M60A1 Loader is:

1. Inspect an M219 machinegun.
2. Stow main gun rounds in tank.

The respondents were informed that if they chose 1 in the example, they would get a Loader who could inspect the machinegun but could not stow main gun rounds. If they chose 2, they would get a Loader who could stow rounds but could not inspect the M219.

---

[1]McCormick, E.J. and Bachus, J.A. Paired comparison ratings. I. The effect on ratings of reductions in the number of pairs. Journal of Applied Psychology, April, 1952.

Table 2

NUMBERS OF PAIRS OF TASKS IN EACH OF THE TWELVE FORMS OF THE PAIRED COMPARISON QUESTIONNAIRE

Tank \ Crew Pos.	Driver	Loader	Gunner	Tank Commander
M60A1	245	231	135	135
M48A5	280	266	135	141
M60A3	252	195	189	171

Each respondent's questionnaire dealt with only one crew position and only one tank. The respondents completed their questionnaires at home, and were encouraged to call a member of the project staff if questions arose.

Additional details about the instructions to the respondents may be found in Appendix E.

RESULTS

Criticality values were calculated for each of the twelve sets of tasks by a standard three step procedure.[1] First, the number of times a task was chosen by the respondents was converted to a proportion by dividing by the number of times it could have been chosen. The number of times a task could have been chosen was the product of the number of respondents (three or four)[2] and the number of pairings for the task (six or seven). The proportions were then changed to normal deviates, *z*. Finally, the *z* values within each task set were transformed to standard scores with a mean of 5.00 and standard deviation of 1.00. This final transformation placed the 12 sets of values on a similar positive scale.

Criticality values of the tasks are shown by tank and duty position in Appendix B. Tasks representative of the high and low ends of the criticality scale are shown in Figure 1, where it can be seen that the top rated tasks are those that would be expected by one familiar with tank operations: the Tank Commander acquiring targets, the Tank Commander or Gunner firing the main gun, the Loader loading, and the Driver driving tactically.

---

[1]Guilford, J.P. op. cit., 1954.

[2]Three Captains did not return their questionnaires.

CREW POSITION	CRITICALITY	TASK
Tank Commander	High	. Acquire Ground Targets (night) . TC Fires Main Gun Precision Using RFD (BEEHIVE) . Zero Tank Main Gun
	Low	. Boresight Searchlight Using Alternate Method (XENON) . Troubleshoot M2 Machinegun . Remove Periscope M36E1 Head Assembly
Gunner	High	. Fire Main Gun Precision Using TEL (Sta/Mov) . Immediate Action In Case of Main Gun Failure to Fire . Performs Main Gun Prepare-To-Fire Procedures
	Low	. Position Gun Tube In Cradle In Response To Signals . Place Turret Into Manual Operation . TC Fires Nonprecision .50 Caliber Using TPI (Sta/Mov)
Loader	High	. Perform Emergency Closing of Main Gun Breech . Load Tank Main Gun Perform Main Gun Prepare-To-Fire Procedures (Loader's Station)
	Low	. Perform Before-Operations Checks On Air Cleaners . Remove M37 Periscope . Check Track Tension
Driver	High	. Perform Evasive Maneuvers On Enemy Contact . Move Vehicle Into Defilade On Enemy Contact . Perform Before-Operations Checks On Engine And Transmission
	Low	. TC Fires Nonprecision Coax Using RFI (Sta/Mov) . Place Turret Into Power Operation . Perform After-Operations Checks On Fender And Stowage Boxes

Figure 1. Tasks representing the extremes in criticality ratings.

Inter-rater reliability was estimated by correlating scale values for tasks common to the three tanks. For example, 27 of the 113 Loader tasks are performed by Loaders on both the M60A1 and the M60A3; the two independently obtained sets of scale values for these 27 tasks were correlated. Correlations, computed by crew position in this manner for each pair of tanks, are shown in Table 3. They ranged from .55 to .79, with an average of .68. All were statistically significant ($p < .05$).

Table 3

RELIABILITY OF CRITICALITY RATINGS
FOR TASKS COMMON TO PAIRS OF TANKS

Crew Position \ Tank Pair	M60A1 M48A5 (N)[1]	M60A1 M60A3 (N)	M48A5 M60A3 (N)	AVG[2]
Commander	.69 (32)	.59 (16)	.79 (7)	.70
Gunner	.71 (35)	.72 (17)	.71 (12)	.72
Loader	.55 (61)	.65 (27)	.64 (25)	.62
Driver	.74 (41)	.64 (44)	.65 (27)	.68

[1](N) = Number of tasks common to the pair of tanks.

[2]AVG = Means based on Fisher's z_r transformation, from Snedecor, G.W. and Cochran, W.G. Statistical Methods (Sixth Edition). Ames, Iowa: Iowa State University Press, 1967.

DISCUSSION

The criticality ratings and inter-rater reliability raise separate issues for discussion, as do questions about the validity of the results obtained.

Criticality

The tasks that were rated high in criticality make sense from a rational or intuitive point of view. Tank Commanders acquiring targets, Gunners firing the main gun, Loaders loading, and Drivers driving tactically, all seem essential for effective performance in combat. But the low-rated tasks -- Check Track Tension, for example, and Place Turret in Manual Operation -- present some interpretive difficulty. The raters' judgments may have been influenced by the likelihood that another crewman could perform the task if the designated crewman could not, or that the task would not have to be performed during a combat mission. Recall also that all the rated tasks had been designated in earlier studies as critical or important.

Reliability

The reliability of the criticality data, though statistically significant and probably greater than the reliabilities of criticality ratings in studies using absolute ratings,[1] seems only marginally acceptable in a practical sense: With a mean inter-rater reliability of .68, the common variance is only about 50 percent. Considering the size of the training investments that are made to teach tasks whose criticality is established by methods less rigorous than the one used here, a search for ways to increase the reliability of criticality ratings seems warranted. Comparing characteristics of the present study with characteristics of other studies may be instructive. No studies other than Boldovici et al.[2] could be found

---

[1]See for example, Harris, J.H., Campbell, R.C., Osborn, W.C., and Boldovici, J.A. Development Of A Model Job Performance Test For A Combat Occupational Specialty. Volume 1. Test Development. Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1975.

[2]Boldovici, J.A., Wheaton, G.R., and Boycan, G.G., op. cit., 1975.

in which reliabilities of criticality estimates higher than those obtained here were reported. The earlier study differed from the present one in several important respects.

The dimensions on which judgments were made were more obvious in the earlier study than in the present one. Target type and target range were the only dimensions along which items were varied in the earlier study. In the present study, the dimensions along which criticality judgments were to be made were less clear. Respondents were simply asked to choose who they would want to take into combat, based on tasks that could or could not be performed by the chosen crew member. The obvious difficulty here is that the nature of the combat or the mission was not specified as clearly as it could have been. Respondents were told only that the mission would involve exchanging fire with the enemy. Given such a vague set, respondents could and undoubtedly did "make up" missions, which differed from one respondent to another. Depending on the anticipated mission, one could, for example, just as easily justify choosing a Loader who could stow main gun rounds as choosing a Loader who could inspect an M219 machinegun. If the respondent doing the ratings was thinking of a recon-by-fire mission or encountering soft targets hidden in a cane field, his choice of a Loader would be different from the choice of a respondent who was thinking of tank-to-tank combat.

The earlier study, in contrast to the present one, left little room for subjects' "making up" the dimensions along which their judgments of criticality would be made. Given a choice, for example, between engaging a tank at 500 meters or a light-armored vehicle at 2500 meters, the dimensions for making the choice are clear:

1. Which target is closer? and
2. Which target is more likely to be equipped with the ammunition, and other means for killing me?

The tank at 500 meters wins on both counts. More importantly, given the absence of opportunity for engaging both targets simultaneously, few if any tankers would disagree with the decision to engage the tank at 500 meters before engaging the light-armored vehicle at 2500 meters. This leads to a second salient difference between the present and the earlier study.

Subjects in the earlier study had certain learning experiences in common, which contributed substantially to high agreement about which one of two targets to engage first: As noted earlier, Tank Commanders receive formal instruction in assessing target threat. The high inter-rater reliability, therefore, may be viewed simply as an index of the extent to which all Tank Commanders had learned the "same things."

Another important difference is that the earlier study, while it did not use complete pairings, more closely approximated a complete pairing design than did the present study. To the extent that complete pairings eliminate the "*luck of the draw*" in determining which tasks get paired with one another, inter-rater reliability would be expected to increase with increases in the number of possible pairs. Some support for this hypothesis is suggested in the literature,[1,2,3,4] though the studies cited differed in many important respects from the present one; in the number of raters, for example, in the total number of stimulus items, in numbers of ratings per pair of items, and in kinds of dependent variables.

---

[1]McCormick, E.J. and Bachus, J.A., op. cit., 1952.

[2]McCormick, E.G. and Roberts, W.K. Paired comparison ratings. 2. The reliability of ratings based on partial pairings. Journal of Applied Psychology, 1952.

[3]Rambo, W.W. Paired comparison scale value variability as function of partial pairing, Psychological Reports, 1959.

[4]Rambo, W.W. The effects of partial pairing on scale values derived from the method of paired comparisons, Journal of Applied Psychology, 1959.

Finally, each stimulus ("task") was rated by more judges in the earlier study than in the present study. To the extent that increasing the number of judges per stimulus decreases systematic bias in the ratings, inter-rater reliability would be expected to increase with increases in the number of judges.

Validity

The conduct of this or any other study that purports to measure task criticality raises questions about the validity of the results obtained, namely:

1. Construct validity: To what extent has what has been purported to have been measured (that is, task criticality) actually been measured? Or, to what extent has inadvertent measurement of constructs other than criticality affected the results obtained?
2. Content validity: To what extent do the "items" (tasks) used in the questionnaires represent the universe of items or tasks?
3. Predictive validity: To what extent would the criticality scores or predictions made from them, correlate with a direct measure of criticality?

Construct Validity. The instructions to the raters in the present study were intended to create a set for judging criticality and criticality alone. But the extent to which the subjects' judgments were influenced by extraneous considerations such as learning difficulty, performance difficulty, performance frequency, and the like is unknown. Questions about construct validity will remain as long as reasonable counterinterpretations of the results can be advanced.[1] Construct validity cannot therefore be established by conducting a "one-shot" study. A plan for initiating examination of

---

[1]Cronbach, L.J. Test validation. In R.L. Thorndike, (Ed.) Educational Measurement (Second Edition), Washington, D.C.: American Council on Education, 1976.

the construct validity of criticality as measured here is presented in Appendix F. The plan is for a correlational study of validity, based on the work of Campbell and Fiske.[1] Factors that might be expected to compete with or contaminate the criticality construct are each measured by two dissimilar methods, as is criticality. The underlying assumption is that measures of the same constructs by dissimilar methods should converge, while measures of different constructs by the same or different methods should diverge.

Content Validity. The issue of how well the content of the questionnaire sampled the universe of subject matter about which conclusions were drawn can never be fully resolved. Resolution would require widespread agreement on the adequacy of the parameters or descriptors used to define the universe, and on precise definition of what constitutes adequate sampling. In the present study, the "universe" was defined as consisting of all tasks rated critical or important in earlier studies by the Army and its contractors; and tasks were sampled from the universe for inclusion in the questionnaires using the method described in Appendix D. To the extent that other investigators would define the task universe differently than was done here, would sample tasks differently, or both, the question of content validity remains open.

As is the case for construct validity, investigation of content validity is not a "one-shot" affair. A duplicate-construction experiment[2] would provide a rigorous test of content validity: Two teams of equally competent questionnaire developers independently would prepare the questionnaires using identical universe definitions

[1]Campbell, D.T. and Fiske, D.W. Convergent and discriminant validation by the multitrait multimethod matrix. Psychological Bulletin, 56, 1959.

[2]Cronbach, L.J., op. cit., 1976.

and rules for selecting questionnaire items. If the universe and sampling are adequately defined, the two forms of the questionnaire will be equivalent. The results of an individual's taking both forms should be identical (within the limits of sampling error).

> "A favorable result, on a suitable broad sample of persons, would strongly suggest that the test content is fully defined by the...construction rules.... An unfavorable result would indicate that the universe definition is too vague or too incomplete to provide a content interpretation for the test."[1]

A less rigorous examination of content validity might be made using critical incidents gathered from veterans of armored combat. Incidents could be gathered until, on the basis of increasing redundancy or another criterion, one was satisfied that the universe of incidents had been adequately sampled. An attempt would then be made to match each task used in the questionnaires with at least one incident. If incidents were identified for which there was no matching task, a basis would be provided for questioning the content validity of the questionnaires. (If, on the other hand, tasks were identified for which there were no matching critical incidents, this would indicate that the pool of critical incidents did not constitute an adequate sample of the task universe.)

Predictive Validity. Establishing the predictive validity of the results of the criticality study would require correlating the obtained criticality scores with a direct measure of criticality. Obtaining direct measures of task criticality in combat is, of course, out of the question. "Direct" is, however, a relative term. Intermediate criteria -- combat simulations, for example -- might be used in studies of predictive validity. One suspects, though, that

---

[1] Cronbach, L.J., op. cit., 1976.

achieving adequate measurement reliability under simulated combat conditions would be very expensive (though absolutely essential if any important decisions are to be made based on the simulation results). Until reliable intermediate criterion measures are forthcoming, the door to establishing the predictive validity of criticality ratings will remain closed.

The more general question of how well indirect measures (ratings, for example) of criticality predict more direct measures may, however, be answerable. Assume, for example, that one could create a game with a clearly defined goal, and with clearly defined tasks that may be performed in achieving that goal. Assume further that, by virtue of design, the relevance or criticality of each task is known to the game's creators. People could be taught the rudiments of the game, given practice until they were thoroughly familiar with its play, and then asked to judge criticality of the various tasks in play of the game. The correlation between task ratings and actual criticality would offer evidence as to the quality of subjective measures of task criticality typically made for real jobs. This hypothetical game could also provide a setting for studying the quality of ratings as a function of job (game) proficiency and rating method.

CONCLUSIONS

1. The criticality values obtained in this study seem to make sense -- more so for the high-rated tasks than for the low-rated tasks. The study, however, dealt only with tasks that had been rated critical or important in earlier studies. Because this was so, and because the present study generated relative criticality ratings, an unavoidable outcome was that some tasks judged critical in earlier studies were judged less critical in the present one.

2. The reliability of the criticality ratings is acceptable, if only marginally so. The paired comparison technique holds promise, and additional research would shed light on how to generate criticality estimates that were highly reliable. Until such research is forthcoming, some tentative operating assumptions can be offered. Inter-rater reliability in studies of task criticality can be expected to increase with:

A. Specificity of the dimensions along which criticality ratings are to be made. This probably is the sine qua non for high rater agreement. To the extent that investigators can create a uniform set among raters as to the dimensions along which judgments are to be made, rater agreement should increase. Without clear specification of the dimensions for making judgments, raters will "make up" their own dimensions. And if these dimensions differ from one rater to the next, rater agreement will suffer.

B. Common learning experiences among raters. The obvious recommendation -- that raters should practice making judgments of the kind required by the criticality study -- is warranted only when the condition discussed in item 1, above is met; that is, when the dimensions for making the judgments are clearly specified. Practice might otherwise simply reinforce idiosyncratic rater behavior and thus reduce rater agreement.

C. The extent to which complete pairings of the tasks to be rated is approximated. The desirability of eliminating the "luck of the draw" in determining which tasks get paired with one another must, however, be traded off against the heavy subject workloads that characterize complete pairings with large numbers of stimulus materials.

D. <u>The number of times each stimulus is rated.</u> Every subject need not rate every possible pair of tasks, though this may be desirable. Decreasing the workload of each subject can be accomplished in several ways. Partial pairings can be used, with all subjects rating all pairs. Or complete pairings can be used with some of the subjects rating some pairs and not others. Various mixes of the approaches also may be used -- partial pairings, with some subjects rating some pairs and not others. The optimal compromises are, unfortunately, not known. Examinations would be interesting, of the effects of various reductions (combined and in isolation) in number or proportion of compared pairs, number or proportion of subjects rating each pair, and number of observations per stimulus and pair on rater agreement. The generality of the results of such research would, of course, never be fully established. Questions would always remain about the effects of stimulus materials, instructions to raters, rater experience, and so forth, on the results obtained. But if confidence is desired in the results of studies that purport to measure the criticality of combat tasks, then additional research on factors affecting rater reliability seems necessary.

The paired comparison method, in any event, would seem to yield reliability estimates that are higher than those found in more conventional ratings of task criticality. But to be more certain, controlled studies comparing various rating methods are needed, especially since inter-rater reliability of criticality ratings is not customarily reported in Army training development literature.

3. The validity of the task criticality ratings remains unknown. Construct, content, and predictive validity present separate issues for consideration:

   A. A plan for initiating investigations of construct validity has been presented. Implementing the plan would shed light on the issue of the extent to which the present study measured criticality, as opposed to other constructs.

B. The issue of content validity never is fully resolved. Suggestions were made, however, for appropriate examinations.

C. No direct measures of the criticality of combat tasks can be made, and intermediate criteria — combat simulations, for example — are likely to be unreliable. Until reliable intermediate criterion measures are forthcoming, the door to establishing predictive validity will remain closed. An approach was suggested, however, for addressing the general question of how well indirect measures of criticality predict more direct measures.

Concern with the validity of the ratings, though appropriate, seems premature. Reliability issues associated with estimating the criticality of armor tasks have only begun to be raised. Given a) that nothing is known about the validity of criticality estimation, and b) choices between results of known and unknown reliability; training developers would seem well advised to use results whose reliability is known.

CLUSTER ANALYSIS

With tasks generated and organized for the three tank systems, and task criticality established with an acceptable degree of reliability, attention was turned to exploring new treatments of the task data. An attempt would be made to identify relatively homogeneous families of tasks, and to use the families as a basis for designing instructional modules in Task 2 of the project.

Cluster analysis[1,2] is a method for sorting or classifying objects, concepts, tasks, or other "things" by measuring similarities among patterns of descriptors. All objects or tasks to be sorted are first described, binary-fashion (yes-no, present-absent), in terms of a common set of descriptors. A simple example of the binary method of description is shown in Figure 2, where three tanks have been characterized according to a common set of descriptors. A cluster analysis of the one-zero data in Figure 2 would sort the tanks by measuring the similarities among the patterns of descriptors that characterize the tanks. The M48A5 and the M60A1 would form a cluster, because their descriptor patterns (1, 0, 0, 1) are identical. The M60A3 would form a separate cluster, because its descriptor pattern (1, 1, 1, 1) is different from the patterns for the M48A5 and the M60A1.[3]

---

[1] Hartigan, J.A., op. cit., 1972.

[2] Dixon, W.J., op. cit., 1975.

[3] The formation of clusters is not as automatic as described here. The process is, in fact, amalgamative and comprised of successive "passes" through the data. In the first pass, each described object forms a cluster. Successive passes form fewer and fewer clusters, each containing more and more of the described objects, until in the final pass, all objects are included in a single cluster. Selecting passes and clusters from the available ones requires devising and using guidelines or rules which reflect the purpose of the analysis. This point is elaborated in Appendix L.

	four-man crew	stabilization	laser rangefinder	105mm gun
M48A5	1	0	0	1
M60A1	1	0	0	1
M60A3	1	1	1	1

Figure 2. Example of one-zero data of the kind used in cluster analysis.

Statistical formulations obviously are not necessary for sorting such disparate objects as tanks. Cluster analysis has, however, been used to study such diverse topics as neighborhood voting preferences,[1] psychosis and anxiety,[2] and tank gunnery job objectives.[3] Cluster analysis was selected for use in the present study in an attempt to identify "families" of armor tasks that had many descriptors in common. If relatively homogeneous families of tasks could be identified, the families could be treated as skills, and efficiency might be achieved in training by designing instructional modules around the skills.

PURPOSE

The main purpose of this part of the project was to examine the utility of cluster analysis as a method for sorting armor tasks. As in the criticality study, the issue of inter-rater reliability also arises: given identical descriptors, tasks, and instructions, to what extent will raters agree on their characterizations of the tasks? A secondary purpose was therefore to examine the extent of correspondence between two independently generated sets of one-zero task description data.

---

[1]Tryon, R.C. Identification of social areas by cluster analysis, University of California, Publications in Psychology, 30, 1955.

[2]Tryon, R.C. Unrestricted cluster and factor analysis with applications to the MMPI and Holtzinger-Harman problems, Multivariate Behavioral Research, 1, 1966.

[3]Boldovici, J.A., Wheaton, G.R., and Boycan, G.G., op. cit., 1976.

METHOD

The method for generating the required one-zero task description data was comprised of two steps:

1. Selecting task descriptors.
2. Characterizing the tasks.

Selecting Task Descriptors

Several criteria were used in selecting descriptors for characterizing the tasks. The three main criteria were that:

1. Characterizing the tasks in terms of the descriptors could be done with a reasonable degree of rater agreement. This was seen as the minimal test of the replicability of the procedures used here. The desire to meet the requirement for reasonable inter-rater reliability in turn suggested other criteria for selecting the descriptors; namely, that the descriptors should be definable in ways that would be readily and uniformly understood by the raters. Ideally, the descriptors would be mutually exclusive, though this was recognized at the outset to be a criterion that never would be fully met.
2. Sorting the tasks in terms of similarities among their descriptor patterns should yield differential implications for training. Application of the criterion led, as will be seen later, to considering using existing learning and task taxonomies as descriptors.
3. The descriptors should be comprehensive: All tasks for the three tanks should be describable in terms of the same set of descriptors. Comprehensiveness may, of course, be achieved by the use of a single non-discriminating descriptor for all tasks; "performed by a tank crew member," for example. This consideration led to a final loose criterion concerning number and kind of descriptors, which was applied in conjunction with the comprehensiveness criterion: The descriptors were to be neither so numerous as to be unmanageable nor so few as to mask important distinctions among the tasks.

Consideration was given during early project planning to using the job- task-elements in the Position Analysis Questionnaire[1] as task descriptors. Any job or task, including the tank crew jobs and tasks addressed in this project, almost certainly can be described using the P.A.Q. elements. But cluster analysis based on tasks characterized by the P.A.Q. descriptors would have no clear implications for training. Attention was therefore directed toward finding a set of descriptors which had training principles or learning algorithms associated with it. The obvious candidates were the conditions and kinds of learning described by Gagné,[2] and by Gagné and Briggs[3]; and the learning algorithms presented in the Training Analysis and Evaluation Group's (TAEG) <u>A Technique for Choosing Cost-Effective Instructional Delivery Systems</u>.[4]

Gagné's types of learning were not used. Even though learning principles are presented for each, the eight types of learning are hierarchically ordered, so that any given type may subsume other types that are lower in the hierarchy. The types of learning therefore are not at all mutually exclusive, and this was thought to invite poor discrimination in the task characterizations that would be performed later.

The TAEG's twelve learning types seemed "less hierarchical" than Gagné's, but here again unreliability in task ratings seemed to be invited by the algorithms' not being mutually exclusive. Many tasks and subtasks can be imagined, for example, that one rater would call "Rule Learning and Using," that another rater would call "Making Decisions,"

---

[1]McCormick, E.J., Mecham, R.C., and Jeanneret, P.R. <u>Position Analysis Questionnaire (PAQ)</u>. West Lafayette, Indiana: PAQ Services, Inc., 1972.

[2]Gagné, R.M., <u>op</u>. <u>cit</u>., 1965.

[3]Gagné, R.M., and Briggs, L.J. <u>Principles of Instructional Design</u>. New York, New York: Holt, Rinehart and Winston, Inc., 1974.

[4]Braby, R., Henry, J.M., Parrish, W.F., Jr., and Swope, W.M. <u>A Technique for Choosing Cost-effective Instructional Delivery Systems (TAEG Report No. 16)</u>. Orlando, Florida: Department of the Navy, Training Analysis and Evaluation Group, 1975.

and that yet another would call both. In reviewing the TAEG reports we also noticed that the training guidelines associated with each of the twelve kinds of learning were highly similar. Thus if the TAEG system were used, one might end with no clear-cut implications for differentially applying the guidelines to each kind of learning.[1]

Reviewing the systems discussed above prompted the thought that using a set of descriptors comprised of four subsets might produce results that had differential implications for training:

1. A Stimuli subset, which would allow noting for each task and subtask the cues that initiated and maintained performance. Describing tasks in terms of the stimulus subset would, it was hoped, provide clues later for specifying or selecting training and testing materials, and for specifying display characteristics for training devices.

2. A subset of Tools, Instruments and Controls, which would allow noting for each task and subtask the manipulanda or mediators of crew members' performance. As with the stimulus subset, it was hoped that describing tasks in terms of the tools, instruments, and controls would facilitate selecting training and testing materials, and specifying training device characteristics.

3. A Mediating Processes subset, which would allow noting for each task and subtask the kinds of learning involved in task performance. Most of the TAEG learning classes could be used in this subset, in the interest of providing a fall-back position in the event that clustering tasks on the basis of all four subsets of descriptors would not yield obvious training implications.

4. An Overt Response subset, which would allow noting, for each task and subtask, the motor behavior involved in task performance. Describing tasks in terms of the Overt Response subset would, it was hoped, help in specifying

[1]This is by no means an indictment of the TAEG system. The best training methods or principles for various kinds of learning may well be more similar than different. And there is certainly no reason to believe that types of learning should be or are mutually exclusive. The point is simply that without mutual exclusivity, inter-rater reliability in task classification probably will suffer.

control characteristics of devices, and in test development.

As can be inferred from the foregoing discussion, the criterion of mutual exclusivity (and therefore inter-judge agreement) was "traded off" in the Mediating Process subset against the apparent desirability of using the TAEG descriptors, for which learning algorithms were readily available. The four subsets of descriptors that were selected for use in the study were an amalgam of the TAEG classes of learning, and several stimulus, tool, test equipment, and response descriptors that were included for the sake of definitional clarity, comprehensiveness, or both. The four subsets of descriptors are listed across the top of Figure 3. Definitions of the descriptors are attached as Appendix G.

Characterizing the Tasks

Forms were printed which had the four subsets of task descriptors across the top of the page, and tasks and subtasks down the left side. Figure 3 is a part of one of the forms. Generating the task by descriptor matrix began with selecting 18 of the 226 M60A1 tasks for use in practicing the task characterizations or ratings. Two criteria were used in selecting the 18 practice tasks:

1. Each duty position was represented in the sample in approximately the same proportion as the duty position is represented in the population of M60A1 tasks.
2. The sample tasks represented the types of tasks performed by each crew member. The Driver was represented by maintenance and driving tasks, for example, and the Gunner by coax and main gun tasks.

Two members of the project staff independently rated the subtasks for each of the 18 sample tasks. Working from left to right in the row corresponding to each subtask (see Figure 3), each rater entered a "1" in the columns corresponding to descriptors that characterized the subtask, and left blank the descriptor columns that did not pertain to the subtask.

DRIVER

TASK No. AD101

	STIMULI											TLS, INSTMTS CONTRLS						MEDIATING PROCESSES											OVERT RESPONSES							
	1. Written (textual) material	2. Graphic/tabular material	3. Instrument read-outs	4. Natural environmental features	5. Man-made environmental features	6. Oral command or request	7. Non-verbal sounds	8. Smell (olfaction)	9. Body feel (kinesthesis)	10. Touch	11. Self-initiated	12. Cmn hnd tls and measuring devices	13. Spec hnd tls and measuring devices	14. On-off or open-close cntrls	15. Fixed setting controls	16. Variable setting controls	17. None	18. Recalls bodies of knowledge	19. Uses verbal information	20. Uses rules	21. Makes decisions	22. Detects (vigilance)	23. Classifies	24. Identifies symbols	25. Recalls set procedures	26. Estimates speed	27. Estimates distance	28. Adopts proper attitude	29. Finger manipulation	30. Hand-arm movement	31. Foot-leg movement	32. Steers	33. Tracks	34. Reports in writing	35. Reports by talking	36. None
PERFORM BEFORE-OPERATIONS CHECKS ON HYDRAULIC BRAKE SYSTEM																																				
• Apply brake and hold for approximately 30 seconds.																																				
• Observe brake pressure gage and insure that it indicates and maintains 750-900 PSI.																																				
• Note any drop in pressure as a fault on DA Form 2404.																																				

Figure 3. Part of the data matrix corresponding to one task.

The ratings were done at the subtask rather than the task level in the interest of inter-rater reliability: Assuming that greater precision is possible in defining subtasks than in defining tasks, one would expect the reliability of the ratings to be greater at the subtask than at the task level.

The raters based their judgments on their knowledge of the conditions under which the subtasks are normally performed, the behavior involved in performing the subtasks, information from technical manuals for the vehicles, and the definitions of the task descriptors shown in Appendix G.

On completing the practice ratings, the raters discussed points of disagreement and made notes that increased the clarity and precision of the definitions of the task descriptors. All tasks for each duty position in each of the three tanks were then rated for record independently by the two raters. Note that in performing this final round of ratings, the judges re-rated the 18 tasks that they had rated earlier.

After all subtasks in a given task were rated, each descriptor column was examined. If at least one "1" was noted in the column, then a "1" was entered in same descriptor column for the task. The one-zero entries in the task rows of the two raters' data sheets were used to examine inter-rater reliability. The two raters later reconciled any differences between their data sheets, producing a uniform set of one-zero data which were the input for the cluster analyses.

ANALYSES AND RESULTS

Two kinds of analyses were done using the data generated by the two raters:

1. Inter-rater reliability analyses, to determine:

   A. The extent of agreement between the two raters in characterizing the tasks.

   B. Whether the discussions between the raters after rating the 18 practice tasks improved agreement on their ratings for record.

2. Cluster analyses, to identify skills, or clusters of tasks with descriptor patterns that were dissimilar among clusters and similar within clusters.

Inter-rater Reliability

The extent of agreement between the two raters was studied in two stages. The first stage used the ratings of the 18 practice tasks mentioned earlier. Recall that the 18 practice tasks were interspersed among 226 M60A1 tasks and were rated for record after the practice session by the same two raters who did the practice ratings. Two sets of ratings were therefore available for the 18 practice tasks: the practice ratings, and the ratings for record that were done a month after the practice ratings. Recall also that between the practice ratings and the ratings for record the raters discussed points of disagreement and revised the definitions of the task descriptors for increased precision and clarity. A basis was thus provided for examining the effects of the raters' discussion on inter-rater reliability.

The second stage of the inter-rater reliability study provided an estimate of the final level of reliability achieved. After all tasks were rated, 22 of the 208 M60A1 tasks that were not rated in the practice session were selected using the same criteria as were used for selecting the 18 practice tasks. The ratings for the 22-task sample were compared with the second round of ratings for the 18-task sample, as a means of verifying the level of inter-rater reliability attained in the final round of ratings for the 18 practice tasks, and of checking on the independence of the final ratings of the 18 practice tasks. The tasks comprising the two samples are presented in Appendixes H and I.

Inter-rater reliability was estimated conservatively, using a method that did not count a zero-zero match between raters as an agreement. Phi coefficients (ϕ) were used in all cases as the index of inter-rater reliability. Details of computation, and discussions of the results are presented in Appendix J.

Inter-rater reliability for the 18 tasks rated before discussion was .58, and after discussion .72. The increase was significant at the .05 level.[1] Overall inter-rater reliabilities for all tasks rated after practice were about .70. This is far in excess of chance expectancy, and marginally acceptable in a practical sense. Suggestions for improving inter-rater reliability in studies of this kind are presented in Appendix J.

Task Clusters

The reconciled one-zero task by descriptor data were analyzed using a canned cluster analysis program.[2] The program uses the Direct Clustering algorithm, which is discussed further in Appendix L.

Eight cluster analyses were performed:

1. Across duty positions, M60A1.
2. Across duty positions, M48A5.
3. Across duty positions, M60A3.
4. Across duty positions, across tanks.
5. Driver, across tanks.
6. Loader, across tanks.
7. Gunner, across tanks.
8. Tank Commander, across tanks.

---

[1]The difference was evaluated statistically using a chi-square type analysis of the transformed Fisher's z correlation (Hays, 1967, p. 532).

[2]Dixon, W.J., op. cit., 1975.

The results of the first four analyses were not particularly instructive.[1] The remaining four will be addressed here. The reason for focusing on the last four of the analyses is threefold:

1. The alternative, analyzing the results by tank across duty position was not particularly useful from a training-development point of view, since training normally is done by duty position.
2. Tasks that are more similar within than among tanks should form unique clusters in the analyses by duty position across tanks.
3. The analyses by duty position across tanks should reveal areas and degrees of task similarity across tanks.

The clusters or "skills" for each duty position, their titles,[2] and the tasks comprising each are shown in Appendix B. Eighty skills were identified -- 21 for the Driver, 19 for the Loader, 20 for the Gunner, and 20 for the Tank Commander. Notice that several of the skills (Driver's Clusters 2, 5, 8, 9, and 21, for example) are one- or two-tank clusters. This suggests that unique skills were not masked by the across-tank, by duty-position cluster solutions.

The cluster titles and the descriptor patterns that characterized each skill are shown by duty position in Figures 4, 5, 6, and 7. In each figure, "X" indicates that the descriptor appeared in more than 50 percent of a cluster's tasks, and "/" indicates that the descriptor appeared in 30 to 50 percent of a cluster's tasks. An asterisk after a cluster title indicates that the cluster is comprised of tasks that are functionally dissimilar. Lubricate Machineguns (Loader's Cluster 12), for example, contains the task, "Install Main Gun Breechblock" (see Appendix B). The occasional quirks in cluster composition probably came about because some of the descriptors were not sufficiently "fine-grained" to permit discrimination among some functionally dissimilar tasks; that is,

---

[1]Presented under separate cover to the ARI/Fort Knox Field Unit Chief.

[2]How cluster titles were derived is discussed in Appendix K.

	NO. OF TASKS IN CLUSTER	1. Written (textual) material	2. Graphic/tabular material	3. Instrument read-outs	4. Natural environmental features	5. Man-made environmental features	6. Oral command or request	7. Non-verbal sounds	8. Smell (olfaction)	9. Body feel (kinesthesis)	10. Touch	11. Self-initiated	12. Cmn hnd tls and measuring devices	13. Spec hnd tls and measuring devices	14. On-off or open-close cntrls	15. Fixed setting controls	16. Variable setting controls	17. None	18. Recalls bodies of knowledge	19. Uses verbal information	20. Uses rules	21. Makes decisions	22. Detects (vigilance)	23. Classifies	24. Identifies symbols	25. Recalls set procedures	26. Estimates speed	27. Estimates distance	28. Adopts proper attitude	29. Finger manipulation	30. Hand-arm movement	31. Foot-leg movement	32. Steers	33. Tracks	34. Reports in writing	35. Reports by talking	36. None
		STIMULI											TLS, INSTMTS CONTRLS						MEDIATING PROCESSES											OVERT RESPONSES							
1. INSTALL AND REMOVE EQUIPMENT	13											X	/		X											X					X						
2. DRAIN WATER FROM FUEL FILTER (AVDS 1790-2A ENGINE)	1		X									X	X	X	X	X								X		X					X						
3. MAINTAIN BASIC ISSUE ITEMS	1	X	X									X			X									X							X					X	
4. INSTALL IR PERISCOPE	3	X			X						X				X	X										X				X	X						
5. PERFORM AFTER-OPERATIONS MAINTENANCE ON FUEL SYSTEM AND DRAIN VALVES	2	X	X												X	/							X								X						
6. PERFORM MISCELLANEOUS MAINTENANCE OPERATIONS	19	X	X								/	X	X		/								X	/		X				X	X				X		
7. FILL OUT FORMS	3			X								X	X													X				X					X		
8. PERFORM AFTER-OPERATIONS CHECKS ON DRAIN VALVES	1		X													X							X			X					X						
9. DISCONNECT TRACK	1						X						X	X												X					X						
10. PERFORM TANK OPERATIONS PROCEDURES	7														X	X	X														X	X					
11. PLACE IR PERISCOPES INTO OPERATION	3	X			X							X			X	X	X							X		X					X						
12. ACQUIRE TARGETS	1				X	X		X				X							X				X	X		X		X	X							X	
13. MAINTAIN DRIVER'S INSTRUMENTS AND CONTROLS	11	X	X	/								X	X		X	X	X						X			X				X	X	/			X		
14. ADJUST TRACK TENSION	2			X							/	X	X	X					X				X	X		X					X						
15. PERFORM AFTER-OPERATIONS MAINTENANCE ON AIR CLEANERS	1		X				X					X	X	X									X	X		X				X	X				X	X	
16. DRIVE TACTICALLY*	39				X	X	X					X					X		X			X		X							X	X	X				
17. PREPARE TANK FOR CROSS COUNTRY TOW	1	X	X				X						X	X	X	X	X		X	X						X											
18. MAINTAIN SUSPENSION SYSTEM	6	X	X								X	X	X	X					X				X	X		X				X	X				X		
19. PERFORM AFTER-OPERATIONS MAINTENANCE ON TRACK TENSION	1	X	X										X			X	X		X												X	X	X			X	
20. START TANK ENGINE*	11			X	/		X	X		/	X	X		/	X	X	X		X				X	/		X					X	X				X	
21. MONITOR INSTRUMENT DISPLAYS	4	X	X	X			X					X	X		X	X	X		X				X	X		X				X	X	X			X	X	

Figure 4. Descriptor patterns for Driver clusters.

	NO. OF TASKS IN CLUSTER	1. Written (textual) material	2. Graphic/tabular material	3. Instrument read-outs	4. Natural environmental features	5. Man-made environmental features	6. Oral command or request	7. Non-verbal sounds	8. Smell (olfaction)	9. Body feel (kinesthesis)	10. Touch	11. Self-initiated	12. Cmn hnd tls and measuring devices	13. Spec hnd tls and measuring devices	14. On-off or open-close cntrls	15. Fixed setting controls	16. Variable setting controls	17. None	18. Recalls bodies of knowledge	19. Uses verbal information	20. Uses rules	21. Makes decisions	22. Detects (vigilance)	23. Classifies	24. Identifies symbols	25. Recalls set procedures	26. Estimates speed	27. Estimates distance	28. Adopts proper attitude	29. Finger manipulation	30. Hand-arm movement	31. Foot-leg movement	32. Steers	33. Tracks	34. Reports in writing	35. Reports by talking	36. None
		STIMULI											TLS, INSTMTS CONTRLS						MEDIATING PROCESSES											OVERT RESPONSES							
1. PERFORM TACTICAL LOADING	16						X					X			X	X	X		X							X				X	X					X	
2. PERFORM TACTICAL SAFE-TO-FIRE PROCEDURES	18						X					X				X															X					X	
3. GROUND GUIDE A TANK	4				X	X	X					X							X					X							X						
4. PREPARE TANK RADIO FOR OPERATION	1			X								X			X	X			X							X					X						
5. BORESIGHT MACHINEGUNS	2					X	X					X	X		X	X			X					X		X					X					X	
6. DISASSEMBLE AND REMOVE MACHINEGUNS*	10						/					X	/		X											X				/	X						
7. PERFORM MISFIRE/IMMEDIATE ACTION PROCEDURES	4						X				/	X		X	X	X							/			X				X	X					X	
8. CONDUCT SUSPENSION SYSTEM CHECKS	2	X	X				/				X	X	/	X					X							/				/	X				/	X	
9. TROUBLESHOOT MACHINEGUNS	2	X	X				X								X	X	X		X					X						X	X						
10. OPERATE TANK INTERCOM	1						X								X	X	X		X					X							X					X	
11. PREPARE MISCELLANEOUS TANK COMPONENTS FOR OPERATION*	3						X				/		X		X	X								X						/	X			/			
12. LUBRICATE MACHINEGUNS*	3							X			X	X	X	X	X	X			X				/	X		X				X	X						
13. PREPARE CVC HELMET FOR OPERATION	1							X			X	X	X						X				X	X		X				X	X						
14. PERFORM MAIN GUN PREPARE-TO-FIRE PROCEDURES	1			X			X	X			X	X	X	X	X	X	X		X				X	X		X				X	X					X	
15. PERFORM MAINTENANCE CHECKS AND SERVICES*	19	/	/				/				/	X	X						/				X			/					X					X	
16. PLACE GUN TUBE IN TRAVEL LOCK	1						X								X		X									X					X					X	
17. BORESIGHT OPTICS	3					X	X							X										X												X	
18. ASSEMBLE/INSTALL MACHINEGUNS	13							/			X	X	X		X	/			/				X			X				X	X					/	
19. OPERATIONAL CHECKS*	9	/	/								/	X	/			/							/			X				X	X				/		

Figure 5. Descriptor patterns for Loader clusters.

	NO. OF TASKS IN CLUSTER	STIMULI											TLS, INSTMTS CONTRLS					
		1. Written (textual) material	2. Graphic/tabular material	3. Instrument read-outs	4. Natural environmental features	5. Man-made environmental features	6. Oral command or request	7. Non-verbal sounds	8. Smell (olfaction)	9. Body feel (kinesthesis)	10. Touch	11. Self-initiated	12. Can hnd tls and measuring devices	13. Spec hnd tls and measuring devices	14. On-off or open-close cntrls	15. Fixed setting controls	16. Variable setting controls	17. None
1. ENGAGE TARGETS*	28					X	X								X	X	X	
2. PERFORM PREPARE TO FIRE PROCEDURES	3	X				X	X						X		X	X	X	
3. BORESIGHT SPECIAL SIGHTS	2					X						X				X	X	
4. PREPARE RANGE CARDS	2			X	X	X	X									X	X	
5A. OPERATE TURRET	4				/	/	X						X		X	X	X	
5B. PERFORM MISFIRE PROCEDURES	2							X		X		X			X	X	X	
6. ASSIST IN RANGECARD ENGAGEMENT	4			X	X		X								X	X	X	
7. CONDUCT FIRE-CONTROL INSTRUMENT CHECKOUT*	10			X							/	X			X	X	X	
8. BORESIGHT SEARCHLIGHT*	5			X		X	X										X	
9. ASSIST IN NIGHT .50 CALIBER ENGAGEMENT*	3			/		X									X		X	
10. OPERATE ELEVATION AND GUNNER'S QUADRANT	2			X								X		/			X	
11. PERFORM ZERO PRESSURE CHECKS	1	X		X								X	X		X		X	
12. PERFORM COMPUTER ELEVATION CHANNEL CHECK	1		X	X								X	X		X	X	X	
13. BORESIGHT MACHINEGUNS	2					X					X	X	X		X		X	
14. PREPARE AZIMUTH INDICATOR	1			X		X						X			X		X	
15. ASSIST IN TARGET ENGAGEMENTS*	19				X		X								X	X		
16. DRAIN REPLENISHER SYSTEM	1						X						X		X			
17. INSTALL/TEST SIGHTING SYSTEMS*	4	/	/	/		/					/	X	/		X	X		
18. PREPARE TANK FOR BORESIGHT	1			X		X	X						X				X	
19. FILL REPLENISHER	2	X		X			X				X		X	X			/	
20. PERFORM CHECKS AND SERVICES ON PERISCOPE	2	X	X														/	

	MEDIATING PROCESSES											OVERT RESPONSES							
	18. Recalls bodies of knowledge	19. Uses verbal information	20. Uses rules	21. Makes decisions	22. Detects (vigilance)	23. Classifies	24. Identifies symbols	25. Recalls set procedures	26. Estimates speed	27. Estimates distance	28. Adopts proper attitude	29. Finger manipulation	30. Hand-arm movement	31. Foot-leg movement	32. Steers	33. Tracks	34. Reports in writing	35. Reports by talking	36. None
1. ENGAGE TARGETS*	X				X	X		X					X		X	X		X	
2. PERFORM PREPARE TO FIRE PROCEDURES	X				X	X		X				X	X		X		X	X	
3. BORESIGHT SPECIAL SIGHTS					X	X		X				X	X		X				
4. PREPARE RANGE CARDS					X	X		X				X	X		X			X	
5A. OPERATE TURRET					X			X					X					X	
5B. PERFORM MISFIRE PROCEDURES								X					X	X				X	
6. ASSIST IN RANGECARD ENGAGEMENT	X					X		X					X	X				X	
7. CONDUCT FIRE-CONTROL INSTRUMENT CHECKOUT*	X				X	X		X					X					/	
8. BORESIGHT SEARCHLIGHT*	X					X		X				X	X		X				
9. ASSIST IN NIGHT .50 CALIBER ENGAGEMENT*					X			X					X		X			X	
10. OPERATE ELEVATION AND GUNNER'S QUADRANT	/		/			X		X				X	X		X		/	/	
11. PERFORM ZERO PRESSURE CHECKS	X				X	X		X					X						
12. PERFORM COMPUTER ELEVATION CHANNEL CHECK	X							X					X			/	X	X	
13. BORESIGHT MACHINEGUNS	X					X		X				X	X		X				
14. PREPARE AZIMUTH INDICATOR					X	X		X					X		X			X	
15. ASSIST IN TARGET ENGAGEMENTS*													X					X	
16. DRAIN REPLENISHER SYSTEM								X					X						
17. INSTALL/TEST SIGHTING SYSTEMS*					X			X					X						
18. PREPARE TANK FOR BORESIGHT						X		X					X		X				
19. FILL REPLENISHER	X							X					X						
20. PERFORM CHECKS AND SERVICES ON PERISCOPE					/	/		/					/						

Figure 6. Descriptor patterns for Gunner clusters.

Cluster	NO. OF TASKS IN CLUSTER	STIMULI											TLS, INSTMTS CONTRLS						MEDIATING PROCESSES											OVERT RESPONSES							
		1. Written (textual) material	2. Graphic/tabular material	3. Instrument read-outs	4. Natural environmental features	5. Man-made environmental features	6. Oral command or request	7. Non-verbal sounds	8. Smell (olfaction)	9. Body feel (kinesthesis)	10. Touch	11. Self-initiated	12. Cmn hnd tls and measuring devices	13. Spec hnd tls and measuring devices	14. On-off or open-close cntrls	15. Fixed setting controls	16. Variable setting controls	17. None	18. Recalls bodies of knowledge	19. Uses verbal information	20. Uses rules	21. Makes decisions	22. Detects (vigilance)	23. Classifies	24. Identifies symbols	25. Recalls set procedures	26. Estimates speed	27. Estimates distance	28. Adopts proper attitude	29. Finger manipulation	30. Hand-arm movement	31. Foot-leg movement	32. Steers	33. Tracks	34. Reports in writing	35. Reports by talking	36. None
1. OPERATE WEAPON SYSTEMS	20											X			X	X	X		X				X	X		X				X	X						
2. ADJUST HEADSPACE AND TIMING	2							X				X		X		X	X		X				X			X					X						
3. INSTALL AND REMOVE EQUIPMENT	10											X			X	X	X									X					X						
4. PERFORM TARGET RANGE INPUT (LASER)	1			X		X						X			X	X	X		X				X			X					X						
5. PERFORM MAIN GUN PREPARE TO FIRE PROCEDURES	2			X								X	X	X	X	X	X		X				X	X		X				X	X					X	
6. PERFORM TACTICAL GUNNERY PROCEDURES	28			X		X	X					X			X	/	X		X			X	X	X		X				X	X		X			X	
7. TROUBLESHOOT MACHINEGUNS	2	X	X					X							X	X			X							X				X	X						
8. ASSEMBLE AN M2 MACHINEGUN	1							X			X	X	X		X		X		X				X			X				X	X						
9A. BORESIGHT AND ZERO WEAPONS	9			X		X						X	/		X	X	X		X					X		X				X	X		X			X	
9B. FIRE RANGECARD ENGAGEMENT	2		X	X		X	X					X			X		X		X			X	X	X		X				X	X		X	X		X	
10. OPERATE TANK RADIO	2						X								X		X		X												X					X	
11. ASSIST IN RANGECARD ENGAGEMENTS	4		X	X			X					X			X		X					X				X				X	X					X	
12. ILLUMINATE TARGETS	1		X			X	X					X			X	X	X		X				X			X					X		X			X	
13. PREPARE RANGECARDS	2		X	X	X	X	X					X	X		X	X	X						X	X		X				X	X		X		X	X	
14. BORESIGHT SEARCHLIGHT	1				X							X	X		X	X			X					X		X					X					X	
15. ACQUIRE TARGETS	1				X	X		X				X		X					X	X	X		X	X		X		X								X	
16. OPERATE SEARCHLIGHT	1				X						X	X				X	X						X	X		X					X					X	
17. PREPARE OPTICAL EQUIPMENT FOR OPERATION	2										X	X	/		X	X								X							X					X	
18. ACTIVATE SMOKE GRENADE LAUNCHER	1					X						X			X	X						X	X	X							X						
19. INSTALL AND MAINTAIN OPTICAL EQUIPMENT*	5	/	/	/	/							X			/		/					X				/					X						
20. SERVICE MACHINEGUNS	2										X	X	X						X												X						

Figure 7. Descriptor patterns for Tank Commander clusters.

some descriptors (natural and environmental features, for example) were so broad that tasks that were quite dissimilar operationally could have had identical or very similar descriptor patterns. The fact that this happened as seldom as it did is encouraging: the tasks comprising each cluster do, on the whole, seem to "go together" operationally or functionally.

Narrative descriptions of a sample of the skills and a few representative tasks are shown in Figures 8, 9, 10 and 11. How the narratives were formed is discussed in Appendix L.

The results of the cluster analysis revealed some task clusters that were unique to a particular vehicle, and yielded cluster profiles that enable comparisons among skills for the different duty positions. More generally the results suggested that, in terms of the descriptors used, there tends to be greater similarity across vehicles in tasks performed than there is between functional categories of tasks within a vehicle. In other words, tasks representing similar tank operations tended to cluster together regardless of which tank they are performed on.

One can, in retrospect, think of several ways that the descriptors could be changed for more desirable cluster definitions. Task complexity or difficulty is not reflected in the descriptors as well as it could have been; for example, the stimulus descriptor "man-made environmental features," would be checked in one instance for a white panel boresight target, and in another instance for an obscured tank target to be identified and fired on with the main gun. Or a "variable control" could in one case refer to a dial to be set, and in another case to the Gunner's tracking control handle.

Some of the characteristics that separated the clusters probably are not as important as others for training development purposes; on-off controls, versus fixed setting controls, for example. And one can think of some descriptors that probably should have been added; for example, a descriptor or descriptors that separated reactive or highly time-constrained tasks from those that are not. But selecting the "best" set of descriptors

DRIVER CLUSTER 1: INSTALL AND REMOVE EQUIPMENT
Performs fixed procedure hand-arm manipulation of on-off or open-close controls and sometimes common hand tools in voluntary response to scheduled operations.

Sample Tasks:
- Install the M27 periscope.
- Remove the VVS2 Driver's viewer.

DRIVER CLUSTER 16: DRIVE TACTICALLY
Performs continuous steering and multilimb manipulation of variable controls in voluntary response to oral commands and environmental features by recalling facts, making decisions, and classifying information.

Sample Tasks:
- Perform evasive maneuvers upon enemy contact.
- Move vehicle into defilade firing position upon enemy contact.

Figure 8. Sample Driver clusters, narrative descriptions, and representative tasks.

LOADER CLUSTER 7: PERFORM MISFIRE/IMMEDIATE ACTION PROCEDURES
Performs fixed procedure finger-hand-arm manipulation of special tools and on-off and fixed setting controls in response to oral command and sometimes touch by detecting information.

Sample Tasks:
- Apply immediate action to reduce a stoppage of the M219 machinegun.
- Unload misfired main gun round.

LOADER CLUSTER 15: PERFORM MAINTENANCE CHECKS AND SERVICES
Performs fixed procedure hand-arm manipulation of common tools in response sometimes to either oral command or written technical guidance and touch by detecting and sometimes recalling information. Reports orally.

Sample Tasks:
- Perform at-halt checks on engine and transmission oil levels.
- Perform after-operations checks on final drives.

Figure 9. Sample Loader clusters, narrative descriptions, and representative tasks.

GUNNER CLUSTER 1: ENGAGE TARGETS

Performs continuous, sometimes compensatory, and fixed procedure finger-hand-arm manipulation of various controls in response to an oral command and to man-made environmental features by detecting, recalling, and classifying information while communicating orally.

Sample Tasks:

. Gunner fires main gun battlesight engagement using the GPD (stationary/moving).

. Gunner fires main gun precision engagement using the TEL (stationary/moving).

GUNNER CLUSTER 7: CONDUCT FIRE CONTROL INSTRUMENT CHECKOUT

Performs fixed procedure hand-arm manipulation of various controls in voluntary response to instrument readouts and sometimes to touch by detecting, recalling, and classifying information; sometimes reports orally.

Sample Tasks:

. Place ballistic computer into operation.

. Perform Laser Rangefinder (LRF) malfunction detection test.

Figure 10. Sample Gunner clusters, narrative descriptions, and representative tasks.

TANK COMMANDER CLUSTER 6: PERFORM TACTICAL GUNNERY PROCEDURES

Communicates orally and performs continuous steering and fixed procedure finger-hand-arm manipulation of on-off or open-close controls, variable setting controls, and sometimes fixed setting controls in voluntary response to man-made environmental features, and instrument read-outs, by recalling facts, making decisions, detecting, and classifying information.

Sample Tasks:

. TC fires main gun battlesight engagement using the RFD (stationary/stationary).

. TC fires caliber .50 engagement using the TPI (stationary/moving).

TANK COMMANDER CLUSTER 19: INSTALL AND MAINTAIN OPTICAL EQUIPMENT

Performs hand-arm manipulation of on-off controls or variable setting controls in voluntary response to scheduled operations, written technical guidance, instrument read-outs, or natural environmental features by detecting information and sometimes recalling set procedures.

Sample Tasks:

. Install periscope M36E1 head assembly.

. Perform after-operations maintenance checks and services on periscope M36E1.

Figure 11. Sample Tank Commander clusters, narrative descriptions, and representative tasks.

on an a priori basis probably is not possible. The test of the adequacy of the cluster solution used here will be in the utility of the results for designing training in Task 2.

CONCLUSIONS

1. The results of inter-rater reliability studies with two judges characterizing armor tasks in terms of 36 descriptors indicated that:

   A. Inter-rater reliability increased significantly with practice and discussion, irrespective of whether the tasks rated for record were the same as or different from the tasks rated for practice.

   B. Overall inter-rater reliabilities for the tasks rated after practice were about .70.

2. Increases in inter-rater reliability greater than those obtained in the present studies probably could have been achieved with:

   A. Increased precision and clarity of the descriptor definitions.

   B. More practice.

   C. More access to operational equipment, as a means of verifying information obtained from technical manuals and experts.

3. Cluster analysis was, with few exceptions, effective in sorting tasks according to common mission operations. Occasional peculiarities in cluster composition occurred, probably because some of the descriptors were not sufficiently "fine-grained" to permit discrimination among some dissimilar tasks. Increased cluster homogeneity might be achieved with the addition of some descriptors that reflect task difficulty or complexity, and others that would separate reactive or highly time-constrained tasks from those that are not.

4. The utility of cluster analysis for training design has only begun to be explored. Several iterations of the kinds of analyses reported here will be required before the most useful set of task descriptors for training development is found. Additional data treatments also should be explored. Cluster analyses based only on stimulus descriptors, for example, might yield more obvious implications for media and device selection than will the results reported here.

SKILL CRITICALITY, LEARNING DIFFICULTY, AND EVALUATION DIFFICULTY

The final part of exploring new treatments of task data was an attempt to determine the criticality, learning difficulty, and evaluation difficulty of each of the task clusters or skills identified earlier.

SKILL CRITICALITY

The criticality of each task cluster was computed as the mean criticality for the tasks in the cluster. The summary values for each cluster are shown in Tables 4 through 7, and in Appendix B. Though informative in a descriptive sense, cluster criticality seems not particularly useful from the standpoint of training development. Criticality is useful chiefly in establishing training priorities; and to the extent that training programs are geared ultimately to tasks, it is task criticality that matters. The integrity of a cluster, in terms of its behavioral characteristics, would not be materially altered by omitting one or two tasks, but its average criticality could be. Having obtained the values by task, however, enables one to calculate the criticality of any configurations of tasks that might comprise a training module.

LEARNING AND EVALUATION DIFFICULTY

Learning difficulty and evaluation difficulty for the domain of tank crew behavior associated with each descriptor were rated by five members of the project staff. The estimates for each descriptor were averaged across raters. Difficulty estimates for each skill or cluster were then made by adding the descriptor scores for the modal descriptor pattern for each task cluster. The sums were converted to standardized scales for learning and evaluation difficulty, each with a mean of 5.0 and standard deviation of 1.0, the same standard scale as was used for the criticality ratings. Additional details of the methodology for estimating learning and evaluation difficulty are presented in Appendix M.

Table 4

SKILL CRITICALITY, LEARNING DIFFICULTY, AND EVALUATION DIFFICULTY: DRIVER

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
1	INSTALL AND REMOVE EQUIPMENT	Performs fixed procedure hand-arm manipulation of on-off or open-close controls and sometimes common hand tools in voluntary response to scheduled operations.	Install the M27 periscope.	13	M60A1 M48A5 M60A3	4.58	3.54	3.75
2	DRAIN WATER FROM FUEL FILTER (AVDS 1790-2A ENGINE)	Performs fixed procedure hand-arm manipulation of common and special hand tools and measuring devices, as well as on-off/open-close controls and fixed setting controls in response to graphic/tabular material by classifying information.	Drain water from engine primary fuel filter and fuel/water separator (AVDS 1790-2A Engine).	1	M48A5	5.83	4.66	4.60
3	MAINTAIN BASIC ISSUE ITEMS	Performs hand-arm manipulation of on-off or open-close controls in voluntary response to written technical guidance by classifying information. Reports by talking.	Perform after-operations maintenance checks and services on basic issue items.	1	M60A1 M48A5 M60A3	4.78	4.11	4.09
4	INSTALL IR PERISCOPE	Performs fixed procedure finger-hand-arm manipulation of on-off/open-close controls or fixed setting controls in response to natural environmental features, written (textual) material, and touch.	Install the M24 (IR) periscope.	3	M60A1 M48A5 M60A3	4.91	4.23	4.15
5	PERFORM AFTER-OPERATIONS MAINTENANCE ON FUEL SYSTEM AND DRAIN VALVES	Performs hand-arm manipulation of on-off/open-close controls and sometimes fixed setting controls in response to written technical guidance by detecting information.	Perform after-operations maintenance checks on the fuel system.	2	M60A1 M60A3	4.52	3.62	3.55
6	PERFORM MISCELLANEOUS MAINTENANCE OPERATIONS	Performs fixed procedure hand-arm manipulation of common hand tools/measuring devices and sometimes on-off/open-close controls in voluntary response to written technical guidance and sometimes touch by detecting and sometimes classifying information. Reports in writing.	Perform before-operations checks on exterior and interior fire extinguisher handles.	19	M60A1 M48A5 M60A3	4.75	5.17	5.29
7	FILL OUT FORMS	Performs fixed procedure finger manipulation of common hand tools in voluntary response to instrument readouts: Reports in writing.	Fill out DA From 2408-1 (Daily).	3	M60A1 M48A5 M60A3	5.31	3.89	4.18
8	PERFORM AFTER-OPERATIONS CHECKS ON DRAIN VALVES	Performs fixed procedure hand-arm manipulation of fixed setting controls in response to graphic/tabular material by detecting information.	Perform after operations checks on drain valves.	1	M48A5	3.93	3.73	3.61

Table 4 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
9	DISCONNECT TRACK	Performs fixed procedure hand-arm manipulation of common and special hand tools/measuring devices in response to an oral command.	Disconnect track.	1	M60A1 M60A3	5.01	3.47	3.44
10	PERFORM TANK OPERATIONS PROCEDURES	Performs fixed procedure multi-limb manipulation of various controls in response to oral commands.	Prepare a tank for combat tow.	7	M60A1 M48A5 M60A3	5.05	3.79	3.79
11	PLACE IR PERISCOPES INTO OPERATION	Performs fixed procedure hand-arm manipulation of various controls in voluntary response to natural environmental features and written (textual) material by classifying information.	Place the M24 (IR) periscope into operation.	3	M60A1 M48A5 M60A3	5.50	4.88	4.80
12	ACQUIRE TARGETS	Communicates orally in voluntary response to environmental features and non-verbal sounds by recalling facts, detecting and classifying information, recalling set procedures, estimating distances and adopting a proper attitude.	Acquire ground targets during daylight.	1	M60A1 M48A5 M60A3	5.51	6.30	6.55
13	MAINTAIN DRIVER'S INSTRUMENTS AND CONTROLS	Performs fixed procedure finger-hand-arm and sometimes multi-limb manipulation of common hand tools/measuring devices and various controls in voluntary response to written technical guidance and sometimes instrument read-outs by detecting information.	Perform before-operations maintenance checks on hydraulic brake system.	11	M60A1 M48A5 M60A3	5.23	5.51	5.59
14	ADJUST TRACK TENSION	Performs fixed procedure hand-arm manipulation of common and special hand tools/measuring devices in voluntary response to instrument read-outs and sometimes touch by recalling facts, detecting and classifying information.	Adjust track tension.	2	M60A1 M48A5 M60A3	5.48	5.36	5.34
15	PERFORM AFTER-OPERATIONS MAINTENANCE ON AIR CLEANERS	Performs fixed procedure finger-hand-arm manipulation of common and special hand tools in voluntary response to graphic/tabular material by detecting and classifying information: Communicates orally and reports in writing.	Perform after-operations maintenance checks and services on the air cleaners.	1	M60A1 M48A5 M60A3	5.52	5.47	5.68

Table 4 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
16	DRIVE TACTICALLY*	Perform continuous steering and multilimb manipulation of variable controls in voluntary response to oral commands and environmental features by recalling facts, making decisions and classifying information.	Operate a tank in neutral steer.	39	M60A1 M48A5 M60A3	4.96	6.01	6.03
17	PREPARE TANK FOR CROSS COUNTRY TOW	Performs fixed procedure hand-arm manipulation of common and special hand tools and various controls in response to an oral command and written technical guidance by recalling facts and using verbal information.	Prepare a tank for cross country tow.	1	M60A1 M48A5 M60A3	5.07	5.21	4.71
18	MAINTAIN SUSPENSION SYSTEM	Performs fixed procedure finger-hand-arm manipulation of common and special hand tools/measuring devices in voluntary response to written technical guidance and touch by recalling facts, detecting and classifying information: Reports in writing.	Perform during-halt-in-operations maintenance checks and services on support roller hubs.	6	M60A1 M48A5 M60A3	5.51	6.08	6.01
19	PERFORM AFTER-OPERATIONS MAINTENANCE ON TRACK TENSION	Perform multilimb manipulation of common hand tools/measuring devices, fixed and variable setting controls in response to written technical guidance by recalling facts: Reports by talking.	Perform after-operations maintenance checks and services on track tension.	1	M60A1 M48A5 M60A3	5.49	4.67	4.42
20	START TANK ENGINE*	Performs fixed procedure multilimb manipulation of various controls and sometimes special hand tools in voluntary response to oral commands, non-verbal sounds, instrument read-outs, touch, and sometimes natural environmental features as well as body feel by recalling facts, detecting, and sometimes classifying information: Reports by talking.	Perform main gun prepare-to-fire procedures from the Driver's positio	11	M60A1 M48A5 M60A3	5.21	6.70	7.09
21	MONITOR INSTRUMENT DISPLAYS	Performs fixed procedure multilimb manipulation of common hand tools/measuring devices and various controls in voluntary response to written technical guidance and instrument read-outs by recalling facts, detecting, and classifying information: Communicates orally and reports in writing.	Perform before-operations checks on engine idle speed.	4	M48A5	4.83	6.80	6.73

Table 5

SKILL CRITICALITY, LEARNING DIFFICULTY, AND EVALUATION DIFFICULTY: LOADER

CLUSTER	TITLE	SKILL DESCRIPTION	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
1	PERFORM TACTICAL LOADING	Performs fixed procedure finger-hand-arm manipulation of various controls in response to oral commands by recalling information: Reports orally.	Gunner fires main gun precision engagement using the TEL (stationary/moving)	16	M60A1 M48A5 M60A3	5.33	4.76	4.89
2	PERFORM TACTICAL SAFE-TO-FIRE PROCEDURES	Performs hand-arm manipulation of fixed setting controls in response to oral commands: Reports orally.	TC fires main gun battlesight engagement using the RFD (moving/stationary)	18	M60A1 M48A5 M60A3	5.11	3.48	3.81
3	GROUND GUIDE A TANK	Performs hand-arm movements in voluntary response to oral command and environmental features by recalling and classifying information.	Ground guide a tank	4	M60A1 M48A5 M60A3	3.77	4.70	4.73
4	PREPARE TANK RADIO FOR OPERATION	Performs fixed procedure hand-arm manipulation of on-off and fixed setting controls in voluntary response to instrument read-outs by recalling information.	Prepare tank radio for operation	1	M60A1 M48A5 M60A3	4.27	4.22	4.15
5	BORESIGHT MACHINEGUNS	Performs fixed procedure hand-arm manipulation of common tools and on-off and fixed setting controls in response to oral command and man-made environmental features by recalling and classifying information: Reports orally.	Boresight an M219 machinegun	2	M60A1 M48A5 M60A3	5.65	5.19	5.16
6	DISASSEMBLE AND REMOVE MACHINEGUNS*	Performs fixed procedure hand-arm and sometimes finger manipulation of on-off controls and sometimes common tools usually as a voluntary response, sometimes on oral command.	Disassemble an M219 machinegun	10	M60A1 M48A5 M60A3	4.94	3.75	3.97
7	PERFORM MISFIRE/IMMEDIATE ACTION PROCEDURES	Performs fixed procedure finger-hand-arm manipulation of special tools and on-off and fixed setting controls in response to oral command and sometimes touch by sometimes detecting information.	Unload misfired main gun round	4	M60A1 M48A5 M60A3	5.92	4.73	5.12
8	CONDUCT SUSPENSION SYSTEM CHECKS	Performs procedural hand-arm and sometimes finger manipulation of special and sometimes common tools in response to written technical guidance, touch and sometimes oral command by recalling information: Reports orally and sometimes in writing.	Perform at-halt temperature checks on compensating idler wheel hubs, support roller hubs final driver hubs and shock absorbers	2	M48A5	4.78	5.00	5.05

Table 5 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
9	TROUBLESHOOT MACHINEGUNS	Performs fixed procedure finger-hand-arm manipulation of various controls in response to oral command and written and graphic material by recalling information.	Troubleshoot an M219 machinegun using Table 3-6, TM 9-235-215-10	2	M60A1 M48A5 M60A3	5.83	4.70	4.28
10	OPERATE TANK INTERCOM	Talks and performs hand-arm manipulation of various controls in response to oral command by recalling and classifying information.	Operate vehicular intercommunications equipment	1	M60A1 M48A5 M60A3	4.47	4.11	3.91
11	PREPARE MISCELLANEOUS TANK COMPONENTS FOR OPERATION*	Performs fixed procedure hand-arm and sometimes finger manipulation of common hand tools and on-off fixed setting controls in response to oral command and sometimes touch.	Load smoke grenade launcher	3	M60A1 M48A5 M60A5	5.25	3.79	3.81
12	LUBRICATE MACHINEGUNS*	Performs fixed procedure finger-hand-arm manipulation of common and special tools, on-off and fixed setting controls in voluntary response to touch and non-verbal sounds by recalling, classifying and sometimes detecting information.	Lubricate an M219 machinegun (disassembled into groups and assemblies)	3	M60A1 M48A5 M60A3	5.76	5.85	5.95
13	PREPARE CVC HELMET FOR OPERATION	Performs fixed procedure finger-hand-arm manipulation of common tools in voluntary response to touch and non-verbal sounds by detecting, recalling and classifying information.	Prepare combat vehicle crewman's helmet for operation	1	M60A1 M48A5 M60A3	4.73	5.61	5.76
14	PERFORM MAIN GUN PREPARE-TO-FIRE PROCEDURES	Performs fixed procedure finger-hand-arm manipulation of common and special tools and various controls in response to oral command, instrument read-outs, non-verbal sounds and touch by recalling, detecting and classifying information: Reports orally.	Perform main gun prepare-to-fire procedures from the Loader's position	1	M60A1 M48A5 M60A3	6.37	6.84	7.00

Table 5 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
15	PERFORM MAINTENANCE CHECKS AND SERVICES*	Performs fixed procedure hand-arm manipulation of common tools in response sometimes to either oral command, or written technical guidance and touch by detecting and sometimes recalling information: Reports orally.	Perform at-halt checks on engine and transmission oil levels	9	M60A1 M48A5 M60A3	4.53	4.57	4.87
16	PLACE GUN TUBE IN TRAVEL LOCK	Performs fixed procedure hand-arm manipulation of on-off and variable setting controls in response to oral command: Reports orally.	Place gun tube in travel lock	1	M60A1 M48A5 M60A3	3.51	3.51	3.52
17	BORESIGHT OPTICS	Uses special tool in response to oral command and man-made environmental features by classifying information: Reports orally.	Boresight gunner's telescope	3	M60A1 M48A5	3.52	3.60	3.43
18	ASSEMBLE/INSTALL MACHINEGUNS	Performs fixed procedure finger-hand-arm manipulation of common tools, on-off controls and sometimes fixed setting controls in voluntary response to touch and sometimes non-verbal sounds by detecting and sometimes recalling information: Reports orally.	Clear an M219 machinegun	13	M60A1 M48A5 M60A3	5.57	5.05	5.45
19	OPERATIONAL CHECKS*	Performs fixed procedure finger-hand-arm manipulation of either common tools or fixed-setting controls in voluntary response sometimes to touch and to written technical guidance by sometimes detecting information: Sometimes reports in writing.	Check operation of an M219 machine-gun	9	M60A1 M48A5 M60A3	4.88	4.43	4.58

Table 6

SKILL CRITICALITY, LEARNING DIFFICULTY, AND EVALUATION DIFFICULTY: GUNNER

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
1	ENGAGE TARGETS*	Performs continuous, sometimes compensatory, and fixed procedure finger-hand-arm manipulation of various controls in response to an oral command and to man-made environmental features by detecting, recalling and classifying information while communicating orally.	Gunner fires main gun battlesight engagement using the GPD ~~(stationary/~~ stationary)	28	M60A1 M48A5 M60A3	5.77	6.08	5.92
2	PERFORM PREPARE TO FIRE PROCEDURES	Performs continuous and fixed procedure finger-hand-arm manipulation of various controls and common tools in response to an oral command, written and graphic material, and to man-made environmental features by detecting, recalling and classifying information; reports orally and in writing.	Perform main gun prepare-to-fire procedures from the Gunner's position	3	M60A1 M48A5 M60A3	6.58	6.36	6.06
3	BORESIGHT SPECIAL SIGHTS	Performs continuous and fixed procedure finger-hand-arm manipulation of fixed and variable setting controls in voluntary response to man-made environmental features by detecting and classifying information.	Boresight M35E1 Gunner's periscope	2	M60A3	5.08	5.61	5.61
4	PREPARE RANGE CARDS	Performs continuous and fixed procedure finger-hand-arm manipulation of fixed setting and variable setting controls in response to oral command, environmental features and instrument read-outs by detecting and classifying information; reports orally.	Prepare a sketch rangecard	2	M60A1 M48A5 M60A3	4.48	5.98	5.81
5A	OPERATE TURRET	Performs fixed procedural hand-arm manipulation of various control and common tools to oral command, sometimes environmental features, by detecting information; reports orally.	Operate gun elevating and turret traversing system in stabilized mode	4	M60A1 M48A5 M60A3	4.84	4.61	4.74
5B	PERFORM MISFIRE PROCEDURES	Performs fixed procedural finger-hand-arm manipulation of various controls in voluntary response to non-verbal sounds and body-feel while communicating orally.	Apply immediate action in case of main gun failure to fire	2	M60A1 M48A5 M60A3	6.37	4.90	5.34
6	ASSIST IN RANGECARD ENGAGEMENT	Performs fixed procedural finger-hand-arm manipulation of various controls in response to oral command, instrument read-outs and natural environmental features by recalling and classifying information while communicating orally.	TC fires main gun rangecard lay to direct fire using the RFD (stationary/stationary) (BEEHIVE)	4	M60A1 M48A5 M60A3	4.32	5.24	5.00

Table 6 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
7	CONDUCT FIRE-CONTROL INSTRUMENT CHECKOUT*	Performs fixed procedure hand-arm manipulation of various controls in voluntary response to instrument readouts and sometimes to touch by detecting, recalling, and classifying information; sometimes reports orally.	Inspect tank thermal sight	10	M60A1 M48A5 M60A3	5.39	5.52	5.47
8	BORESIGHT SEARCHLIGHT*	Performs continuous steering and fixed procedure finger-hand-arm manipulation of variable controls in response to oral command, instrument read-puts and man-made environmental features by recalling and classifying information.	Boresight tank searchlight using primary method	5	M60A1 M48A5 M60A3	4.42	5.31	4.80
9	ASSIST IN NIGHT .50 CALIBER ENGAGEMENT*	Performs continuous and fixed procedure hand-arm manipulation of *on-off and variable controls* in response to man-made environmental features and sometimes to instrument read-outs, by detecting information; reports orally.	TC fires nonprecision .50 caliber *engagement using the TPI* (moving/moving)	3	M60A1 M48A5 M60A3	4.44	4.61	4.56
10	OPERATE ELEVATION AND GUNNER'S QUADRANT	Performs continuous and fixed procedure finger-hand-arm manipulation of variable controls and sometimes special tools in voluntary response to instrument read-outs by classifying information and sometimes recalling information and using rules; sometimes reports orally and in writing.	Operate Gunner's quadrant	2	M60A1 M48A5 M60A3	4.15	5.49	5.30
11	*PERFORM ZERO PRESSURE* CHECKS	Performs fixed procedure hand-arm manipulation of common tools on-off and variable controls in voluntary response to textual material and instrument read-outs by detecting and classifying information.	Perform a zero pressure check (hydraulic power pack)	1	M60A1 M48A5 M60A3	5.16	5.57	5.41

Table 6 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
12	PERFORM COMPUTER ELEVATION CHANNEL CHECK	Performs fixed procedure hand-arm manipulation of common tools and various controls in voluntary response to graphic material and instrument read-outs by recalling information; reports orally and in writing.	Perform XM21 computer elevation channel check	1	M60A3	5.39	5.07	4.96
13	BORESIGHT MACHINEGUNS	Performs continuous steering and fixed procedure finger-hand-arm manipulation of common tools, on-off and variable controls in voluntary response to touch and man-made environmental features by recalling and classifying information.	Boresight an M219 machinegun mounted on a tank	2	M60A1 M48A5 M60A3	4.10	5.90	5.85
14	PREPARE AZIMUTH INDICATOR	Performs continuous steering and fixed procedure hand-arm manipulation of on-off and variable setting controls in voluntary response to man-made environmental features and instrument read-outs by detecting and classifying information; reports by talking.	Prepare azimuth indicator for operation	1	M60A1 M48A5 M60A3	3.62	5.63	5.67
15	ASSIST IN TARGET ENGAGEMENTS*	Performs hand-arm manipulation of on-off and fixed setting controls in response to oral command and sometimes natural environmental features: reports orally.	TC fires main gun battlesight engagement using the RFD (moving/stationary)	19	M60A1 M48A5 M60A3	4.48	3.41	3.59
16	DRAIN REPLENISHER SYSTEM	Performs fixed procedure hand-arm manipulation of common tools and on-off controls in response to oral command.	Drain replenisher system	1	M60A1 M48A5 M60A3	4.22	3.32	3.34
17	INSTALL/TEST SIGHTING SYSTEMS*	Performs fixed procedure hand-arm manipulation of on-off and fixed setting controls and sometimes common tools in voluntary response to either written technical guidance and instrument read-outs, or touch or man-made environmental features by detecting information.	Install tank thermal sight	4	M60A3	4.57	4.81	4.96
18	PREPARE TANK FOR BORESIGHT	Performs continuous steering and fixed procedure hand-arm manipulation of common tools and variable controls in response to oral command, instrument read-outs and man-made environmental features by classifying information	Prepare tank for boresighting	1	M60A1 M48A5 M60A3	5.25	4.88	4.51

Table 6 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
19	FILL REPLENISHER	Performs fixed procedure hand-arm manipulation of common and special tools and sometimes variable controls in response to oral commands, written material, touch, and instrument read-outs by recalling information.	Fill replenisher system	2	M60A1 M48A5 M60A3	4.35	4.71	4.42
20	PERFORM CHECKS AND SERVICES ON PERISCOPE	Sometimes performs fixed procedure hand-arm manipulation of variable setting controls in response to written technical guidance by either classifying or detecting information.	Perform before-operations maintenance checks and services on periscope M35E1	2	M60A3	5.39	3.63	3.30

Table 7

SKILL CRITICALITY, LEARNING DIFFICULTY, AND EVALUATION DIFFICULTY: TANK COMMANDER

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	[illegible]	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
1	OPERATE WEAPON SYSTEMS	Performs fixed procedure, finger-hand-arm manipulation of various controls in voluntary response to man-made environmental features, non-verbal sounds, or touch by recalling facts, detecting or classifying information.	Inspect Tank Commander's periscope M36E1	20	M60A1 M48A5 M60A3	4.95	5.32	5.32
2	ADJUST HEADSPACE AND TIMING	Performs fixed procedure hand-arm manipulation of special hand tools and measuring devices, fixed setting controls and variable setting controls in voluntary response to non-verbal sounds by recalling facts and detecting information.	Adjust headspace on the M2 machinegun	2	M48A5	5.39	5.16	5.27
3	INSTALL AND REMOVE EQUIPMENT	Performs fixed procedure hand-arm manipulation of various controls in voluntary response to scheduled operations.	Disassemble an M85 machinegun	10	M60A1 M48A5 M60A3	4.01	3.88	3.99
4	PERFORM TARGET RANGE INPUT (LASER)	Performs fixed porcedure hand-arm manipulation of various controls in voluntary response to instrument read-outs and man-made environmental features by recalling facts and detecting information.	Perform target range input (laser)	1	M60A3	5.30	5.26	5.20
5	PERFORM MAIN GUN PREPARE TO FIRE PROCEDURES	Communicates orally and performs fixed procedure finger-hand-arm amnipulation of various hand-tools and controls in voluntary response to instrument read-outs by recalling facts, detecting and classifying information.	Perform main gun prepare-to-fire procedures from the Tank Commander's position	2	M60A1 M48A5 M60A3	5.43	6.04	6.03
6	PERFORM TACTICAL GUNNERY PROCEDURES	Communicates orally and performs continuous steering and fixed procedure finger-hand-arm manipulation of on-off or open-close controls, variable setting controls and sometimes fixed setting controls in voluntary response to man-made environmental features, and instrument read-outs, by recalling facts, making decisions, detecting and classifying information.	Tank Commander fires nonprecision .50 caliber engagement using TPI (moving/moving)	28	M60A1 M48A5 M60A3	5.48	6.96	6.84
7	TROUBLESHOOT MACHINEGUNS	Performs fixed procedure, finger-hand-arm manipulation of on-off and fixed setting controls in response to non-verbal sounds and written technical guidance by recalling facts.	Troubleshoot an M85 machinegun TM 9-2350-215-10, Table 3-6	2	M60A1 M48A5 M60A3	4.12	4.54	4.11

Table 7 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
8	ASSEMBLE AN M2 MACHINEGUN	Performs fixed procedure finger-hand-arm manipulation of common hand tools, on-off and variable setting controls, in voluntary response to scheduled operations, non-verbal sounds and touch by recalling facts and detecting information.	Assemble an M2 machinegun	1	M48A5	3.78	5.52	5.83
9A	BORESIGHT AND ZERO WEAPONS	Performs continuous and fixed procedure finger-hand-arm manipulation of various controls and sometimes common hand tools in voluntary response to man-made environmental features, instrument read-outs and sometimes touch by recalling facts and classifying information: Reports by talking.	Zero tank main gun	9	M60A1 M48A5 M60A3	5.61	6.04	5.79
9B	FIRE RANGECARD ENGAGEMENT	Performs continuous and compensatory tracking and fixed procedure finger-hand-arm manipulation of on-off and variable setting controls in voluntary response to an oral command, graphic material, instrument read-outs and man-made environmental features by recalling facts, making decisions, detecting and classifying information: Reports by talking.	Tank Commander fires coax rangecard lay to direct fire using the RFI (stationary/moving)	2	M60A1 M48A5 M60A3	5.27	7.57	7.44
10	OPERATE TANK RADIO	Performs hand-arm manipulation of on-off and variable setting controls in response to oral commands: Reports by talking..	Operate tank radio	2	M60A1 M48A5 M60A3	4.72	3.53	3.52
11	ASSIST IN RANGECARD ENGAGEMENTS	Performs fixed procedure finger-hand-arm manipulation of on-off and variable setting controls in voluntary response to an oral command, graphic material and instrument read-outs by making decisions: Reports by talking.	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary) BEEHIVE	4	M60A1 M48A5 M60A3	4.82	5.15	5.16
12	ILLUMINATE TARGETS	Performs continuous steering and fixed procedure hand-arm manipulation of controls in voluntary response to an oral command, graphic/tabular material and man-made environmental features by recalling facts and detecting information: Reports by talking.	Illuminate targets using tank search-light	1	M60A1 M48A5	4.24	5.93	6.01

Table 7 (Continued)

C.U.	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# of TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
13	PREPARE RANGECARDS	Performs continuous steering and fixed procedure finger-hand-arm manipulation of common hand tools, fixed setting controls and variable setting controls in voluntary response to oral commands, environmental features and instrument read-outs by detecting and classifying information: Reports by talking and in writing.	Prepare a sketch rangecard	2	M60A1 M48A5 M60A3	4.54	6.97	7.02
14	BORESIGHT SEARCHLIGHT	Performs fixed procedure hand-arm manipulation of common hand tools, on-off and fixed setting controls in voluntary response to man-made environmental features by recalling facts and classifying information: Reports by talking.	Boresight tank searchlight using primary method	1	M60A1 M48A5	3.88	5.01	5.03
15	ACQUIRE TARGETS	Reports by talking and uses special measuring devices in voluntary response to environmental features and non-verbal sounds by recalling facts, using verbal information, using rules, detecting information, classifying information, recalling set procedures, and estimating distances.	Acquire ground targets (right)	1	M60A1 M48A5 M60A3	6.39	6.88	6.60
16	OPERATE SEARCHLIGHT	Performs fixed procedure hand-arm manipulation of on-off and fixed setting controls in voluntary response to natural environmental features, and touch by detecting and classifying: Reports by talking.	Place tank searchlight into operation	1	M60A1 M48A5	4.09	5.20	5.54
17	PREPARE OPTICAL EQUIPMENT FOR OPERATION	Performs hand-arm manipulation of on-off and fixed setting controls and sometimes common hand tools in voluntary response to touch by classifying information: Reports by talking.	Prepare tank thermal sight for operation	2	M60A3	3.92	4.15	4.40
18	ACTIVATE SMOKE GRENADE LAUNCHER	Performs hand-arm manipulation of on-off and fixed setting controls in voluntary response to man-made environmental features by making decisions, detecting and classifying information.	Activate smoke grenade launcher.	1	M60A3	4.73	4.93	5.02

Table 7 (Continued)

CLUSTER	TITLE	SKILL DESCRIPTIONS	SAMPLE TASKS	# OF TASKS	TANK	CRITICALITY	LRNG DIFF	EVAL DIFF
19	INSTALL AND MAINTAIN OPTICAL EQUIPMENT*	Performs hand-arm manipulation of on-off controls or variable setting controls in voluntary response to scheduled operations, written technical guidance, instrument read-outs, or natural environmental features by detecting information and sometimes recalling set procedures.	Install periscope M36E1 head assembly	5	M60A1 M48A5 M60A5	4.45	4.36	4.35
20	SERVICE MACHINEGUNS	Performs hand-arm manipulation of common hand tools in voluntary response to touch by recalling facts.	Service an M85 machinegun	2	M60A1 M48A5 M60A5	4.54	3.84	4.08

RESULTS AND DISCUSSION

The learning and evaluation difficulty estimates for each skill are presented in Tables 4 through 7. Inter-rater reliability was estimated by an analysis of variance of the rater by descriptor data matrix.[1] Intraclass correlations were .76 for learning difficulty and .88 for evaluation difficulty, indicating fairly high reliability of the average of the five sets of ratings. (Each coefficient indicates the hypothetical correlation that would obtain between the average ratings for this set of five raters and those from another random sample of five raters.) If it is assumed, however, that the raters differed systematically in their frames of reference for judging the descriptors, then the reported correlations are underestimates of inter-rater reliability. When the data are corrected for differences among rater means, reliabilities of the mean ratings are .85 for learning difficulty, and .89 for evaluation difficulty.

Averages of the learning and evaluation difficulty scale values were computed across the skills in each duty position. These means, presented in Figure 12, indicate that the skills required for the Tank Commander's position are the most difficult for learning and for evaluation, followed by the Gunner, Driver, and Loader on both dimensions. These findings supported the expectations of the relative learning and evaluation difficulties of skills among the four duty positions. Figure 12 also presents tasks representative of those skills which received the highest and lowest difficulty scores in each duty position. The same skills appeared at the extremes of both dimensions in each of the four duty positions.

The results of the learning and evaluation difficulty study seemed in some cases to be at odds with reality. Driver's Cluster 20 "Start tank engine," for example, received an evaluation difficulty rating that

---

[1]Winer, B.J. Statistical Principles in Experimental Design. New York, New York: McGraw-Hill, 1962.

CREW POSITION	CRITICALITY	SKILL	TASK
TANK COMMANDER Mean LD[1] = 5.34 Mean ED[2] = 5.36	HIGH	9B. FIRE RANGECARD ENGAGEMENT	. TC Fires Coax Rangecard Lay To Direct Fire Using The RFI (Sta/Mov)
		13. PREPARE RANGECARDS	. Prepare A Circular Rangecard
	LOW	3. INSTALL AND REMOVE EQUIPMENT	. Remove An M85 Machinegun From A Tank
		10. OPERATE TANK RADIO	. Operate Tank Radio
GUNNER Mean LD = 5.08 Mean ED = 4.98	HIGH	1. ENGAGE TARGETS	. Gunner Fires Main Gun Battlesight Engagement Using The GPD (Mov/Mov)
		2. PERFORM PREPARE-TO-FIRE PROCEDURES	. Perform Main Gun Prepare-To-Fire Checks
	LOW	15. ASSIST IN TARGET ENGAGEMENTS	. TC Fires Main Gun Battlesight Engagement Using the RFD (Mov/Sta)
		20. PERFORM CHECKS AND SERVICES ON PERISCOPE	. Perform Before-Operations Maintenance Checks And Services On Periscope M35E1
LOADER Mean LD = 4.63 Mean ED = 4.71	HIGH	12. LUBRICATE MACHINEGUNS	. Lubricate An M219 Machinegun (disassembled into groups and assemblies)
		14. PERFORM MAIN GUN PREPARE-TO-FIRE PROCEDURES	. Perform Main Gun Prepare-To-Fire Procedures From the Loader's Position
	LOW	16. PLACE GUN TUBE IN TRAVEL LOCK	. Place The Gun Tube In Travel Lock
		17. BORESIGHT OPTICS	. Boresight Gunner's Telescope

Figure 12. Representative skills and tasks at the extremes in learning and evaluation difficulty.

DRIVER Mean LD = 4.92 Mean ED = 4.92	HIGH	20. START TANK ENGINE	. Start Tank Engine By Auxiliary Power -- Slave Start (Using M48A5) For Auxiliary Power
		21. MONITOR INSTRUMENT DISPLAYS	. Performs Before-Operations Maintenance Checks On Tank Instruments, Gages, And Warning Lights (Engine Off)
	LOW	1. INSTALL AND REMOVE EQUIPMENT	. Install The M27 Periscope
		5. PERFORM AFTER-OPERATIONS MAINTENANCE ON FUEL SYSTEM AND DRAIN VALVES	. Perform After-Operations Maintenance Checks On The Fuel System

Figure 12 (Continued). Representative skills and tasks at the extremes in learning and evaluation difficulty.

was more than two standard deviations above the mean. Such apparent abberations probably occured for either or both of two reasons. The first is that the method for computing clustèr difficulty was additive. (Recall that difficulty was computed by summing the difficulty values for descriptors that predominated each cluster.) The sum of the values rather than the mean was used, on the assumption that the greater the number of descriptors required to characterize the cluster, the greater the cluster's complexity, and therefore the greater its difficulty of evaluation and learning. This assumption may have been erroneous.

Another possible reason for the apparent abberations is simply that some of the cluster names do not describe the tasks comprising the cluster very well. This is especially true for the asterisked clusters, which were comprised of tasks related to more than one mission operation, but which were named in terms of only one mission operation. The abberant Driver's Cluster 20 mentioned above is, in fact, one of the asterisked clusters. It is comprised, not only of tasks related to starting the engine, but also of operating a tank across a water obstacle, driving over varied terrain, and performing main gun prepare-to-fire procedures -- tasks that may indeed be extremely difficult to evaluate. Time and other resources unfortunately did not permit exploring other ways of computing cluster difficulty that might have produced results different from those obtained. Summing the descriptor difficulty values for each task, for example, and then averaging the task values within each cluster would be interesting.

As was the case with the criticality ratings, a question can be raised about the extent to which learning difficulty and evaluation difficulty were rated independently of other constructs (criticality, for example). The extent to which learning difficulty and evaluation difficulty are independent of one another also may be of interest. These are, of course, questions of construct validity and could be examined using a plan analogous to the one presented for the criticality ratings (see Appendix F). Construct validity also can be examined, albeit tentatively, by correlating some scores from the present study. The learning and evaluation difficulty estimates for the 32 descriptions were highly correlated ($r = .76$). This may indicate that skills that are difficult to learn also are difficult to evaluate. But the learning and evaluation difficulty values were generated on the basis of scores from the same group of raters. The high correlation may, therefore, be a measurement artifact: The two constructs may have been related in the judgment of the raters, but not in fact.

Other correlations bearing on the issue of construct validity are shown in Table 8. The correlations between learning difficulty and criticality, and between evaluation difficulty and criticality averaged .44. As was the case for the correlation between learning and evaluation difficulty, the correlations may reflect a "real" relationship, or systematic bias in the ratings (or both). The criticality estimates and the difficulty estimates were, however, (a) generated from ratings by two independent sets of judges (Captains and project staff members), and (b) measured differently from one another. This suggests that the constructs are related in fact rather than only in the judgment of the raters. Why criticality and difficulty would be related is not clear. Designers of tank systems may, because of space, hardware, or money limitations, allocate the most critical system functions (detecting and tracking targets, for example) to men rather than machines -- and these critical functions may indeed be the most difficult to learn and evaluate.

CONCLUSIONS

1. The cluster criticality estimates, which were averages of the criticality values for the tasks comprising each cluster, probably will not be as useful in training design as the criticality values for individual tasks will be.
2. The estimates of learning evaluation and difficulty were highly reliable in terms of the stability of the mean ratings obtained.

Table 8

CORRELATIONS (r) BETWEEN CLUSTER CRITICALITY AND LEARNING DIFFICULTY; AND BETWEEN CLUSTER CRITICALITY AND EVALUATION DIFFICULTY

	N	Learning Difficulty and Criticality	Evaluation Difficulty and Criticality
Tank Commander	20	.55*	.48
Gunner	20	.20	.22
Loader	19	.61*	.64
Driver	21	.41	.41
Average		.44	.44

*$p < .05$

3. The results of the learning and difficulty studies were inconclusive. Some of the results seemed at odds with reality. This may have been because of deficiencies in methods for computing difficulty, because some of the clusters were named inappropriately, or both. The results reported here can be verified via additional treatments of the obtained data (computing difficulty values for each task, and averaging task values within each cluster, for example), or by conducting additional research (paired comparison studies of task difficulty, for example).
4. The estimates of learning difficulty and evaluation difficulty were highly correlated. Skills that are difficult to learn may tend to be difficult to evaluate also. The possibility of measurement error remains, however, and may be examined using designs similar to the one presented in Appendix F.
5. The estimates of learning difficulty and evaluation difficulty each correlated on an average of .44 with the criticality estimates. The suggestion was offered that criticality and difficulty may in fact be related because of system design practices that assign more critical and difficult system functions to men rather than to machines.

REFERENCES

Ammerman, H.L. and Pratzner, F.C. Occupational Survey on Auto Mechanics: Task Data from Workers and Supervisors Indicating Job Relevance and Training Criticalness. Columbus, Ohio: Ohio State University, 1975.

Anonymous. Phase II Product Improvements for M60A1/M60A1-P1 Tanks. undated.

Boldovici, J.A., Wheaton, G.R., and Boycan, G.G. Selecting Items For a Tank Gunnery Test. Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

Braby, R., Henry, J.M., Parrish, W.F., Jr., and Swope, W.M. A Technique for Choosing Cost-effective Instructional Delivery Systems (TAEG Report No. 16). Orlando, Florida: Department of the Navy, Training Analysis and Evaluation Group, 1975.

Campbell, D.T. and Fiske, D.W. Convergent and discriminant validation by the multitrait matrix. Psychological Bulletin, 56, 1959.

Chrysler Corporation. Operator's Manual for Tank, Combat, Full Tracked: 105-MM Gun, M60A1E3. Detroit, Michigan: Author, 1974.

Cronbach, L.J. Test validation. In R.L. Thorndike, (Ed.) Educational Measurement (Second Edition), Washington, D.C.: American Council on Education, 1976.

Dixon, W.J., (Ed.). BMDP: Biomedical Computer Programs. Berkeley, California: University of California Press, 1975.

Ford, J.P., Harris, J.H., and Rondiac, P.F. Performance Measures for AIT Armor Crewmen. Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1974.

Gagné, R.M. The Conditions of Learning. New York, New York: Holt, Rinehart and Winston, 1965.

Gagné, R.M. and Briggs, L.J. Principles of Instructional Design. New York, New York: Holt, Rinehart and Winston, 1974.

Gropper, G.L. and Short, J.G. Handbook for Training Development. Pittsburgh, Pennsylvania: American Institutes for Research, 1969.

Guilford, J.P. Psychometric Methods. New York, New York: McGraw-Hill, 1954.

Harris, J.H. Task Lists for M60A1, M60A1(AOS), M48A5, and M60A3 Tanks (Project Memorandum No. 1). Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

Harris, J.H., Campbell, R.C., Osborn, W.C., and Boldovici, J.A. Development Of a Model Job Performance Test For a Combat Occupational Specialty. Volume 1. Test Development. Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1975.

Harris, J.H., O'Brien, R.E., Campbell, R.C., and Ford, J.P. Sub-tasks for the M60A1 Tank System (Memorandum No. 4). Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

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

Hays, W.L. Statistics for Psychologists. New York, New York: Holt, Rinehart and Winston, 1965.

McCluskey, M.R., Jacobs, T.O., and Cleary, F.K. Systems Engineering of Training for Eight Combat Arms MOSs, Alexandria, Virginia: Human Resources Research Organization (HumRRO), 1975.

McCormick, E.J. and Bachus, J.A. Paired comparison ratings. I. The effect on ratings of reductions in the number of pairs. Journal of Applied Psychology, April, 1952.

McCormick, E.J., Mecham, R.C., and Jeanneret, P.R. Position Analysis Questionnaire (PAQ). West Lafayette, Indiana: PAQ Services, Inc., 1972.

McCormick, E.J. and Roberts, W.K. Paired comparison ratings. 2. The reliability of ratings based on partial pairings. Journal of Applied Psychology, 1952.

McKnight, J.A. and Hundt, A.G. Driver Education Task Analyses: The Development of Instructional Objectives. Alexandria, Virginia: Human Resources Research Organization (HumRRO), 1972.

O'Brien, R.E. and Boldovici, J.A. Task Lists for Chrysler XM-1 Prototype (Project Memorandum No. 3). Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1976.

Rambo, W.W. Paired comparison scale value variability as function of partial pairing, Psychological Reports, 1959.

Rambo, W.W. The effects of partial pairing on scale values derived from the method of paired comparisons, Journal of Applied Psychology, 1959.

Schumacher, S.P. and Glasgow, A.Z. Handbook for Designers of Instructional Systems, Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratories, 1973.

Snedecor, G.W. and Cochran, W.G. Statistical Methods (Sixth Edition). Ames, Iowa: Iowa State University Press, 1967.

Texas Instruments Incorporated. Tank Thermal Sight, Preliminary Design Review. Lubbock, Texas: Author, 1976.

Tryon, R.C. Identification of social areas by cluster analysis. University of California, Publications in Psychology, 30, 1955.

Tryon, R.C. Unrestricted cluster and factor analysis with applications to the MMPI and Holtzinger-Harman problems, Multivariate Behavioral Research, 1, 1966.

US Army Armor School. 11E Job Task Data Cards. Fort Knox, Kentucky: Author, 1976.

US Army. Operator's Manual for Tank, Combat, Full Tracked: 105-MM Gun, M48A5, NSN 2350-00-582-5595, TM 9-2350-258-10. Author, 1975.

US Army Transportation School. Interservice Procedures for Instructional Systems Development. Fort Eustis, Virginia: Author, 1975.

Winer, B.J. Statistical Principles in Experimental Design. New York, New York: McGraw-Hill, 1962.

APPENDIXES

APPENDIX A

METHOD FOR GENERATING THE TASK LISTS

METHOD FOR GENERATING THE TASK LISTS

M60A1 TASK LIST

Three data sources were used in generating the M60A1 task and subtask list (see Table 1, p. 7). The main data source for the M60A1 list was a set of job task data cards for the critical and important communications, machinegun, and tracked vehicle tasks, as indicated in the 11E task list, and supplied by the Job and Task Analysis Branch, Directorate of Training Developments, U.S. Army Armor School, Fort Knox, Kentucky (1976). Task data and criticality ratings from the Armor School were supplemented by task data and criticality ratings from a second source, Performance Measures for AIT Armor Crewmen.[1]

Gunnery tasks for the M60A1 list were obtained from a third source. Boldovici, Wheaton, and Boycan[2] attempted to define all tasks encompassed by M60A1(AOS) gunnery.[3] Since the task lists in that study seemed more comprehensive than any available others, they were used to sample gunnery tasks for use in the present project. Two criteria were used for selecting the gunnery tasks -- comprehensiveness and representativeness.

Comprehensiveness refers to the extent to which the gunnery tasks as a group cover the gunnery domain, as represented in Table A.1. Representativeness refers to the extent to which a task in each cell of the domain subsumes elements or subtasks of other tasks in the same cell.

---

[1]Ford, J.P., Harris, J.H., and Rondiac, P.F. Performance Measures for AIT Armor Crewmen. Fort Knox, Kentucky: Human Resources Research Organization (HumRRO), 1974.

[2]Boldovici, J.A., Wheaton, G.R., and Boycan, G.G., op. cit., 1976.

[3]This study updated an earlier attempt at domain definition by Kraemer, Boldovici, and Boycan (1975).

Table A.1

LOCATIONS IN THE GUNNERY DOMAIN, OF TASKS USED IN THIS PROJECT
(Each "X" represents one task.)

WEAPON / FIRE DELIVERY METHOD	MAIN GUN TC	MAIN GUN GNR	COAX TC	COAX GNR	CAL .50 TC
Battlesight (non-precision for machineguns)	X	XX	X	X	X
Precision	X	XX XX	X	X	
Range Card					
Range card Lay to Direct Fire	X	X	X	X	

Preliminary results from the Boldovici, Wheaton, and Boycan[1] study identified those gunnery tasks that were most comprehensive and representative of the M60A1(AOS) gunnery domain. Their locations in the domain are shown in Table A.1. The 17 gunnery tasks were modified to incorporate a stationary firing vehicle, and became part of the M60A1 task list for the present project.[2]

M48A5 TASK LIST

Generating the M48A5 list began with a review of the M60A1 list. All tasks that were rated critical or important for the M60A1 in the sources described earlier, and that would be performed by M48A5 crew members, were considered also to be critical or important for the M48A5 and were included in the M48A5 list. The M60A1-based list for the M48A5 was expanded in two ways:

1. The M48A5 Operator's Manual was reviewed. Whenever a task was found that was performed by an M48A5 crew member, but not by an M60A1 crew member, we made a judgment about the criticality or importance of the task. If it was judged critical or important, the task was added to the M48A5 list.
2. The gunnery tasks that were included in the M48A5 list were the same as the gunnery tasks for the M60A1. They were the set of tasks, modified to incorporate target engagements from a stationary firing vehicle, which according to the Boldovici, Wheaton, and Boycan report were most comprehensive and representative tasks in the M60A1(AOS) gunnery domain.

The M48A5 task list included 22 more tasks than the M60A1 list did. These were tasks which the project staff judged important or critical, but which were not in the 11E most-critical and important lists supplied by the Armor School. Examples of the added tasks included, "Check track tension," "Connect track," and "Zero M2 machinegun."

---

[1]Boldovici, J.A., Wheaton, G.R., and Boycan, G.G., op. cit., 1976.

[2]The M60A1 task and subtask lists have been presented under separate cover. (See Harris, J.H., O'Brien, R.E., Campbell, R.C., and Ford, J.P., 1976.)

M60A3 TASK LIST

The M60A3 will be the production version of the experimental M60A1E3. Because of uncertainty about which product improvements will be incorporated into the M60A3, some guesswork was required in generating the task list for this tank.

As with the M48A5, the task list for the M60A1 was used as a starting point for generating the list for the M60A3. Any M60A1 task that was also performed by an M60A3 crew member, and was rated critical or important for the M60A1, was included in the M60A3 list. Gunnery tasks were the ones designated most comprehensive and representative in the study by Boldovici, Wheaton, and Boycan.[1] And the M60A1E3 Operator's Manual was reviewed to identify tasks which seemed critical or important to the project staff, but had not appeared in the 11E task list.

Best guesses had to be made, based on interviews with authorities at Fort Knox, and on reviews of product improvement literature, about the final configuration of the M60A3. Task lists were then written for the operation and maintenance of those components that seemed most likely to be incorporated into the production M60A3.

The M60A3 task list that evolved was different in several ways from the M60A1 task list:

1. The M60A3 gunnery tasks included precision engagements from moving tanks with no requirement to come to a brief halt before firing.
2. Tasks were written to reflect the following new components, which are likely to replace existing ones or are new to the tank inventory.
   - A. Laser Rangefinder, ANVVG2 (new component).
   - B. Electronic Computer, XM21 (new component).
   - C. Light Amplification Sights, M35E1, M36E1 (new component for Tank Commander, replaces existing periscope for Gunner).

[1] Boldovici, J.A., Wheaton, G.R., and Boycan, G.G., op. cit., 1976.

D. Tank Thermal Sight (new component).

E. Smoke Grenade Launcher (new component).

F. Muzzle Reference System (new component).

G. MAG-58 Coax Machinegun (replaces M219 machinegun).

H. Driver's Viewer, VVS2 (replaces Driver's viewer, M27).

APPENDIX B

TASK CLUSTERS AND CRITICALITY ESTIMATES

DRIVER

TASK NO:	CLUSTER 1: INSTALL AND REMOVE EQUIPMENT*	CRITICALITY M60A1	M48A5	M60A3
AD105	Install the M27 periscope	5.355		4.402
A5111	Install the M27 periscope (spare)		4.348	
A3125	Remove the VVS2 Driver's viewer			4.876
AF105	Remove the M24 (IR) periscope	4.512		
A5112	Remove the M27 periscope		4.076	
AD106	Remove the M27 periscope	3.782		4.876
A3119	Disconnect final drive universal joints			4.402
A5110	Remove the M24 (IR) periscope		5.238	
AA128	Place a tank external phone into operation	4.827	4.052	4.220
AA113	Inspect universal joints	4.636	5.972	4.402
AA127	Perform operator maintenance on tank external phone	4.759	4.916	4.220
A3120	Disconnect transmission input shaft			3.784
A5128	Replace track pads (T142)		4.543	
	CLUSTER CRITICALITY:		4.583	
	CLUSTER 2: DRAIN WATER FROM FUEL FILTER (AVDS 1790-2A ENGINE)			
A5130	Drain water from engine primary fuel filter and fuel/water separator (AVDS 1790-2A engine)		5.830	
	CLUSTER CRITICALITY:		5.830	
	CLUSTER 3: MAINTAIN BASIC ISSUE ITEMS			
AA120	Perform after-operations maintenance checks and services on basic issue items	4.877	4.586	4.876
	CLUSTER CRITICALITY:		4.780	
	CLUSTER 4: INSTALL IR PERISCOPE			
AF103	Install the M24 (IR) periscope	4.636		
A3123	Install the VVS2 Driver's viewer			5.624
A5108	Install the M24 (IR) periscope		4.468	
	CLUSTER CRITICALITY:		4.909	

*Asterisks indicate clusters comprised of some tasks that are "functionally dissimilar"; that is, tasks that pertain to more than one crew function or mission operation. For details of how the clusters were formed, see text and Appendix L.

TASK NO:	CLUSTER 5: PERFORM AFTER-OPERATIONS MAINTENANCE ON FUEL SYSTEM AND DRAIN VALVES	CRITICALITY M60A1	M48A5	M60A3
AD116	Perform after-operations maintenance checks on the fuel system	5.355		3.484
AD120	Perform after-operations maintenance checks and services on the tank drain valves	4.877		4.371
	CLUSTER CRITICALITY:	4.522		

CLUSTER 6: PERFORM MISCELLANEOUS MAINTENANCE OPERATIONS

TASK NO:		M60A1	M48A5	M60A3
A5102	Perform before-operations checks on exterior and interior fire extinguisher handles		4.916	
A5131	Perform after-operations maintenance checks and services on Driver's hatch		4.348	
AD119	Perform after-operations maintenance checks on tank fire extinguishers	5.355		5.096
AD117	Perform after-operations maintenance checks and services on tank hatches	4.877		3.048
AA123	Perform after-operations maintenance checks and services on the gun travel lock	4.106	3.748	1.296
AD102	Perform before-operations maintenance checks on fire extinguishers	5.355		5.570
A5138	Perform after-operations checks on Driver's escape hatch		4.218	
A5135	Perform after-operations checks on the suspension system		6.511	
A5133	Perform after-operation checks on fire extinguisher handles - external		4.369	
A5106	Perform before-operations maintenance checks and services on Driver's hatch		4.586	
AA124	Perform after-operations maintenance checks and services on the engine and transmission	6.637	5.972	5.170
AA119	Perform after-operations maintenance checks and services on fender stowage boxes	4.512	4.076	1.296
AD118	Perform after-operations maintenance checks and services on the tank batteries	5.479		4.876
AA121	Perform after-operations maintenance checks and services on the tank hull	4.996	5.132	4.402
AD103	Perform before-operations maintenance checks and services on hatches	4.382		4.017
A5132	Perform after-operations checks and services on the tank batteries		5.972	
AF101	Perform before-operations maintenance checks on tank instruments, gages and warning lights	5.610		
AF108	Perform after-operations maintenance checks on tank instruments, gages and warning lights	4.877		
A5129	Perform after-operations maintenance checks on the fuel system		7.191	
	CLUSTER CRITICALITY:	4.750		

CLUSTER 7: FILL OUT FORMS

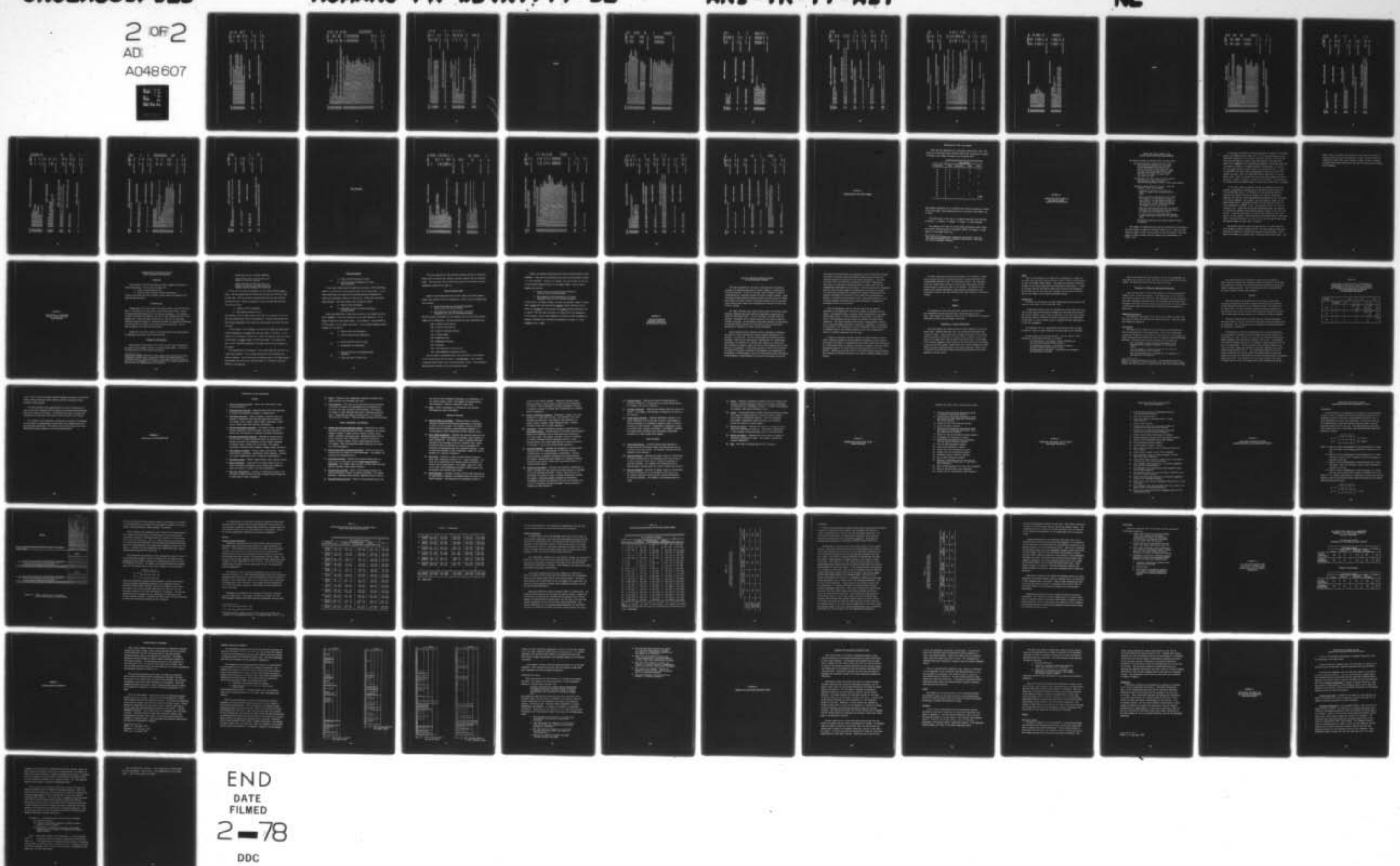
TASK NO:		M60A1	M48A5	M60A3
AA102	Fill out DA Form 2408-1 (Daily)	5.206	5.462	5.170
A5117	Perform during operations checks on instruments, gages and warning lights (engine running)		5.972	
AA103	Fill out DA Form 2404	4.636	4.916	5.821
	CLUSTER CRITICALITY:	5.312		

TASK NO:		CRITICALITY M60A1	M48A5	M60A3
	CLUSTER 8: PERFORM AFTER-OPERATIONS CHECKS ON DRAIN VALVES			
A5134	Perform after-operations checks on drain valves		3.925	
	CLUSTER CRITICALITY:		3.925	
	CLUSTER 9: DISCONNECT TRACK			
AD113	Disconnect track	5.610		4.402
	CLUSTER CRITICALITY:		5.006	
	CLUSTER 10: PERFORM TANK OPERATIONS PROCEDURES			
AA116	Prepare a tank for combat tow	4.877	5.541	4.876
AA111	Stop tank engine	4.636	4.348	4.852
AA104	Perform before-operations maintenance checks and services on tank engine and transmission	6.405	9.720	5.974
AA126	Place turret into power operation	4.106	3.748	3.484
AA114	Prepare a tank for highway tow	4.382	5.024	3.784
AD108	Place a tank in motion	4.996	6.123	
A5121	Prepare an inoperable tank for towing		5.051	
	CLUSTER CRITICALITY:		5.051	
	CLUSTER 11: PLACE IR PERISCOPE INTO OPERATION			
A5109	Place the M24 (IR) periscope into operation		4.810	
AF104	Place the M24 (IR) periscope into operation	5.886		
A3124	Place the VVS2 Driver's viewer into operation			5.791
	CLUSTER CRITICALITY:		5.496	
	CLUSTER 12: ACQUIRE TARGETS			
AA112	Acquire ground targets during daylight	5.206	5.700	5.624
	CLUSTER CRITICALITY:		5.510	

AD-A048 607 HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA F/G 5/9
CRITICALITY AND CLUSTER ANALYSES OF TASKS FOR THE M48A5, M60A1,--ETC(U)
NOV 77 J A BOLDOVICI, J H HARRIS, W C OSBORN DAHC19-76-C-0001

UNCLASSIFIED HUMRRO-FR-WD(KY)77-12 ARI-TR-77-A17 NL

2 OF 2
AD
A048607



END
DATE
FILMED
2-78
DDC

TASK NO:	CLUSTER 13: MAINTAIN DRIVER'S INSTRUMENTS AND CONTROLS	CRITICALITY M60A1	M48A5	M60A3
AD101	Perform before-operations maintenance checks on hydraulic brake system	5.742		5.771
A5101	Perform before-operations maintenance on hydraulic brake system		5.349	
A3128	Perform after-operations maintenance checks on tank instruments, gages, and warning lights			5.624
A3127	Perform before-operations maintenance checks on tank instruments, gages, and warning lights			5.974
A5107	Perform before-operations maintenance checks and services on accelerator and steering controls		4.916	
AB101	Perform before-operations maintenance checks and services on the M24 IR periscope and M27	5.610	5.462	
A5120	Perform before-operations checks on the gas particulate filter unit		4.810	
A3121	Perform before-operations maintenance checks on tank instruments, gages, and warning lights			5.489
A3122	Perform before-operations maintenance checks and services on the VVS2 Driver's IR viewer			5.469
A5119	Perform before-operations checks on personal heater		4.218	
AA122	Perform after-operations maintenance checks and services on tank lights	5.355	4.468	4.220
	CLUSTER CRITICALITY:		5.232	

	CLUSTER 14: ADJUST TRACK TENSION			
AD115	Adjust track tension	5.527		5.317
A5125	Adjust track tension		5.580	
	CLUSTER CRITICALITY:		5.475	

	CLUSTER 15: PERFORM AFTER-OPERATIONS ON AIR CLEANERS			
AA125	Perform after-operations maintenance checks and services on the air cleaners	5.742	5.462	5.341
	CLUSTER CRITICALITY:		5.515	

TASK NO:	CLUSTER 16: DRIVE TACTICALLY*	CRITICALITY M60A1	M48A5	M60A3
AD109	Operate a tank in neutral steer	4.512		4.967
AD114	Connect track	4.996		4.876
AA106	Respond to ground guide signals while driving a tank	5.232	5.024	5.317
A5124	Check track tension		5.700	
A5115	Operate a tank in neutral steer		5.275	
AD111	Operate a tank across a water obstacle	5.206		5.317
AD110	Drive the tank over varied terrain with Driver's hatch in the open/close position	6.934		6.034
A5127	Connect track		5.580	
A5126	Disconnect track		6.123	
AA108	Perform during-operations maintenance checks and services on steering accelerator, shift and brake controls	5.886	5.024	5.317
AD104	Perform before-operations maintenance checks and services on steering, accelerator, transmission and brake controls	6.069		5.408
A3101	Tank Commander fires main gun battlesight engagement using the RFD (moving/stationary)			5.926
AA109	Perform evasive maneuvers upon enemy contact	9.077	6.780	6.709
A3115	Gunner fires main gun precision engagement using the TEL (moving/moving)			5.974
AL117	Gunner fires main gun precision engagement using the GPD (stationary/stationary)(BEEHIVE)	4.636	3.537	
AL116	Gunner fires main gun precision engagement using the GPD (stationary/stationary)	4.382	4.218	
AL115	Gunner fires main gun precision engagement using the TEL (stationary/moving)	4.512	4.348	
AL114	Gunner fires main gun precision engagement using the TEL (stationary/moving)	3.782	3.537	
AL108	TC fires precision coax engagement using the RFD (stationary/stationary)	3.057	3.268	
AL107	TC fires main gun precision engagement using the RFD (stationary/stationary)	4.382	3.367	
AL106	TC fires nonprecision .50 caliber engagement using the TPI (stationary/moving)	3.057	3.748	
AL104	TC fires nonprecision coax engagement using the RFI (stationary/moving)	3.057	3.268	
AL103	Gunner fires main gun battlesight engagement using the GPD (stationary/moving)	4.512	4.348	
AL102	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)	4.250	4.076	
AL101	TC fires main gun battlesight engagement using the RFD (stationary/stationary)	3.782	3.952	
A3117	Gunner fires main gun precision engagement using the GPD (moving/stationary)(BEEHIVE)			5.624
A3116	Gunner fires main gun precision engagement using the GPD (moving/stationary)			6.371
A3114	Gunner fires main gun precision engagement using the TEL (stationary/moving)			4.724
A3109	Gunner fires precision coax engagement using the TEL (moving/stationary)			4.923
A3108	TC fires precision coax engagement using the RFD (stationary/stationary)			4.422
A3107	TC fires main gun precision engagement using the RFD (moving/stationary)(BEEHIVE)			4.750
A3106	TC fires nonprecision .50 caliber engagement using the TPI (moving/moving)			4.876
A3105	Gunner fires nonprecision coax engagement using the TEL (moving/stationary)			5.023
A3104	TC fires nonprecision coax engagement using the RFI (moving/moving)			5.469
A3103	Gunner fires main gun battlesight engagement using the GPD (moving/moving)			6.132
A3102	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)			5.791
AA110	Move vehicle into defilade firing position upon enemy contact	7.372	6.511	7.209
AL105	Gunner fires nonprecision coax engagement using the TEL (stationary/stationary)	3.354	3.748	
AL109	Gunner fires precision coax engagement using the TEL (stationary/stationary)	3.782	3.925	

CLUSTER CRITICALITY: 4.959

	CLUSTER 17: PREPARE TANK FOR CROSS COUNTRY TOW	M60A1	M48A5	M60A3
AA115	Prepare a tank for cross country tow	4.512	6.123	4.565

CLUSTER CRITICALITY: 5.067

TASK NO:	CLUSTER 18: MAINTAIN SUSPENSION SYSTEM	CRITICALITY M60A1	CRITICALITY M48A5	CRITICALITY M60A3
AA107	Perform during-halt-in-operation maintenance checks and services on support roller hubs, road wheel hubs, compensating idler wheel hubs, and final drive hubs	5.282	6.399	5.624
AA113	Inspect universal joints	4.636	5.972	4.402
A3126	Perform after-operations maintenance checks and services on suspension system			5.821
AF107	Perform after-operations maintenance checks and services on suspension system	6.043		
AF102	Perform before-operations maintenance checks and services on suspension system	6.405		
A5104	Perform before-operations maintenance checks and services on suspension system		4.586	
	CLUSTER CRITICALITY:		5.512	

TASK NO:	CLUSTER 19: PERFORM AFTER-OPERATIONS MAINTENANCE ON TRACK TENSION	M60A1	M48A5	M60A3
AA117	Perform after-operations maintenance checks and services on tank track tension	5.206	5.830	5.443
	CLUSTER CRITICALITY:		5.493	

TASK NO:	CLUSTER 20: START TANK ENGINE*	M60A1	M48A5	M60A3
AA105	Perform main gun prepare-to-fire procedures from the Driver's position	4.996	4.699	4.565
A5118	Operate a tank across a water obstacle		5.132	
A5116	Drive the tank over varied terrain with Driver hatch in the open/close position		6.599	
AD112	Start tank engine by towing	4.250		4.402
A5122	Start tank engine by towing		6.300	
A5123	Start tank engine by auxiliary power--slave start (using M48A5) for auxiliary power		5.580	
AD107	Start tank engine	5.479		6.285
A5113	Start tank engine		4.879	
A5114	Place a tank in motion		4.586	
A3118	Start tank engine by auxiliary power-slave start			5.791
AF106	Start tank engine by auxiliary power-slave start	4.636		
	CLUSTER CRITICALITY:		5.212	

TASK NO:	CLUSTER 21: MONITOR INSTRUMENT DISPLAYS	M60A1	M48A5	M60A3
A5105	Perform before-operations checks on engine idle speed		4.586	
A5136	Perform after-operations checks on instruments, gages and warning lights		4.916	
A5103	Perform before-operations maintenance checks on tank instruments, gages and warning lights (engine off)		5.024	
A5137	Perform after-operations checks on engine fuel shut-off switch		4.810	
	CLUSTER CRITICALITY:		4.834	

LOADER

TASK NO:	CLUSTER 1: PERFORM TACTICAL LOADING	CRITICALITY M60A1	M48A5	M60A3
A3214	Gunner fires main gun precision engagement using the TEL (stationary/moving)			5.670
AL217	Gunner fires main gun precision engagement using the GPD (stationary/stationary) (BEEHIVE)	5.812	4.707	
AL211	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary) (BEEHIVE)	5.133	4.745	
AL210	TC fires main gun rangecard lay to direct fire using the RFD (stationary/stationary) (BEEHIVE)	5.542	4.105	
AL207	TC fires main gun precision engagement using the RFD (stationary/stationary) (BEEHIVE)	5.998	5.443	
A3217	Gunner fires main gun precision engagement using the GPD (moving/stationary) (BEEHIVE)			6.071
A3211	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary) (BEEHIVE)			5.670
A3210	TC fires main gun rangecard lay to direct fire using the RFD (stationary/stationary) (BEEHIVE)			5.143
A3207	TC fires main gun precision engagement using the RFD (moving/stationary) (BEEHIVE)			6.071
AK209	Unload a MAG-58 machinegun			5.345
AB211	Unload an M219 machinegun	5.265	5.291	
AL216	Gunner fires main gun precision engagement using the GPD (stationary/stationary)	5.000	4.707	
AL215	Gunner fires main gun precision engagement using the TEL (stationary/moving)	4.735	4.707	
AL214	Gunner fires main gun precision engagement using the TEL (stationary/moving)	5.000	4.999	
A3216	Gunner fires main gun precision engagement using the GPD (moving/moving)			6.071
A3215	Gunner fires main gun precision engagement using the TEL (moving/moving)			6.803

CLUSTER CRITICALITY: 5.334

	CLUSTER 2: PERFORM TACTICAL SAFE-TO-FIRE PROCEDURES	M60A1	M48A5	M60A3
A3201	TC fires main gun battlesight engagement using the RFD (moving/stationary)			6.168
AL213	Gunner fires coax rangecard lay to direct fire using the TEL (stationary/moving)	5.265	4.707	
AL212	TC fires coax rangecard lay to direct fire using the RFI (stationary/moving)	5.116	4.999	
AL209	Gunner fires precision coax engagement using the TEL (stationary/stationary)	5.133	4.707	
AL208	TC fires precision coax engagement using the RFD (stationary/stationary)	6.174	4.078	
AL205	Gunner fires nonprecision coax engagement using the TEL (stationary/stationary)	4.867	4.852	
AL204	TC fires nonprecision coax engagement using the RFI (stationary/moving)	5.403	4.556	
AL203	Gunner fires main gun battlesight engagement using the GPD (stationary/moving)	5.948	4.402	
AL202	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)	5.000	4.556	
AL201	TC fires main gun battlesight engagement using the RFD (stationary/stationary)	5.403	3.706	
A3213	Gunner fires coax rangecard lay to direct fire using the TEL (stationary/moving)			5.345
A3212	TC fires coax rangecard lay to direct fire using the RFI (stationary/moving)			5.143
A3209	Gunner fires precision coax engagement using the TEL (moving/stationary)			4.837
A3208	TC fires precision coax engagement using the RFD (stationary/stationary)			4.942
A3205	Gunner fires nonprecision coax engagement using the TEL (moving/stationary)			5.143
A3204	TC fires nonprecision coax engagement using the RFI (moving/moving)			5.244
A3203	Gunner fires main gun battlesight engagement using the GPD (moving/moving)			6.473
A3202	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)			

CLUSTER CRITICALITY: 5.110

TASK NO:		CRITICALITY M60A1	CRITICALITY M48A5	CRITICALITY M60A3
	CLUSTER 3: GROUND GUIDE A TANK			
AA201	Ground guide a tank	4.458	4.999	2.176
AA228	Disconnect track		4.241	
AA229	Connect track		3.500	
AA230	Check track tension		3.259	
	CLUSTER CRITICALITY:		3.772	
	CLUSTER 4: PREPARE TANK RADIO FOR OPERATION			
AA227	Prepare tank radio for operation	4.189	5.146	3.473
	CLUSTER CRITICALITY:		4.269	
	CLUSTER 5: BORESIGHT MACHINEGUNS			
AB213	Boresight an M219 machinegun	5.998	5.921	
AK211	Boresight a MAG-58 machinegun			5.043
	CLUSTER CRITICALITY:		5.654	
	CLUSTER 6: DISASSEMBLE AND REMOVE MACHINEGUNS*			
AA213	Stow main gun rounds in the tank	5.116	6.292	5.670
AA223	Place gun tube in travel lock	4.597	3.500	3.813
AA222	Perform a zero pressure check (hydraulic power pack)	4.088	4.556	4.216
AA220	Place turret into manual operation	3.640	4.707	5.244
AA221	Place turret into power operation	3.420	4.077	5.244
AB203	Disassemble an M219 machinegun	5.542	6.098	
AA204	Disassemble the breechblock	5.402	5.596	4.496
AK214	Remove a MAG-58 machinegun from a tank			5.143
AB215	Remove an M219 machinegun from a tank	6.360	6.090	
AK201	Disassemble a MAG-58 machinegun			5.647
	CLUSTER CRITICALITY:		6.940	

TASK NO:		CRITICALITY M60A1	M48A5	M60A3
	CLUSTER 7: PERFORM MISFIRE/IMMEDIATE ACTION PROCEDURES			
AD201	Unload misfired main gun round	5.640		5.358
AK212	Apply immediate action to reduce a stoppage of the MAG-58 machinegun			5.143
AB214	Apply immediate action to reduce a stoppage of the M219 machinegun	6.361	6.739	
A5201	Unload misfired main gun		6.098	
	CLUSTER CRITICALITY:		5.923	
	CLUSTER 8: CONDUCT SUSPENSION SYSTEM CHECKS			
A5204	Perform at-halt temperature checks on compensating idler wheel hubs, support roller hubs, final drive hubs and shocks		4.707	
A5211	Perform after-operations maintenance checks-temperature-road wheel, idlers, track support roller hubs		4.852	
	CLUSTER CRITICALITY:		4.780	
	CLUSTER 9: TROUBLESHOOT MACHINEGUNS			
AB212	Troubleshoot an M219 machinegun using Table 3-6, TM 9-235-215-10	6.361	5.443	
AK210	Troubleshoot a MAG-58 machinegun using TABLE 3-6, DEP 9-2350-253-10			5.670
	CLUSTER CRITICALITY:		5.825	
	CLUSTER 10: OPERATE TANK INTERCOM			
AA225	Operate vehicular intercommunications equipment	4.458	4.707	4.258
	CLUSTER CRITICALITY:		4.474	
	CLUSTER 11: PREPARE MISCELLANEOUS TANK COMPONENTS FOR OPERATION*			
AK215	Load smoke grenade launcher			4.363
AB201	Prepare tank for boresighting	5.000	4.707	
AA207	Adjust variable breech operating cam	6.149	5.921	5.345
	CLUSTER CRITICALITY:		5.248	
	CLUSTER 12: LUBRICATE MACHINEGUNS*			
AB206	Lubricate an M219 machinegun(disassembled into groups and assemblies)	5.689	5.443	
AA206	Install main gun breechblock	5.403	6.292	6.071
AK204	Lubricate a MAG-58 machinegun (disassembled into groups and assemblies)			5.670
	CLUSTER CRITICALITY:		5.761	

TASK NO:	CLUSTER 13: PREPARE CVC HELMET FOR OPERATION	CRITICALITY M60A1	M48A5	M60A3
AA224	Prepare combat vehicle crewman's helmet for operation	3.640	5.921	4.617
	CLUSTER CRITICALITY:		4.726	

CLUSTER 14: PERFORM MAIN GUN PREPARE-TO-FIRE PROCEDURES

		M60A1	M48A5	M60A3
AA210	Perform main gun prepare-to-fire procedures from the Loader's position	5.998	7.026	6.071
	CLUSTER CRITICALITY:		6.365	

CLUSTER 15: PERFORM MAINTENANCE CHECKS AND SERVICES*

		M60A1	M48A5	M60A3
A5205	Perform at-halt checks on engine and transmission oil levels		4.611	
A5202	Perform before-operations maintenance checks and services on tank engine and transmission oil levels		4.852	
A5212	Perform after-operations maintenance checks and services on tank engine and transmission oil levels		4.707	
AA218	Prepare a sketch rangecard	3.640	4.707	4.039
AA219	Prepare a circular rangecard	3.420	4.356	3.513
A5207	Prepare an inoperable tank for towing		4.241	
AA212	Perform after-operations maintenance checks and services on tank engine and transmission oil levels	4.528	5.291	5.558
AK202	Inspect a MAG-58 machinegun			5.244
AB204	Inspect an M219 machinegun	5.998	4.707	
AA214	Service tank main gun ammunition	5.000	4.999	5.670
A5206	Perform at-halt checks on final drives		4.295	
A5216	Perform after-operations checks on final drives		4.241	
AB202	Clean and lubricate the breechblock, cannon bore and bore evacuator of the tank after operations	5.000	6.292	
A3202	Clean and lubricate the breechblock, cannon bore and bore evacuator of the tank after operations			5.143
A5213	Perform after-operations maintenance checks and services on the air cleaners and blowers		3.900	
AA226	Perform operator maintenance on radios and accessories	5.000	4.241	4.216
AD202	Perform before-operations maintenance checks and services on tank engine and transmission oil levels	2.442		3.564
A3219	Perform after-operations maintenance checks and services on tank engine and transmission oil levels			3.813
AF201	Perform after-operations maintenance checks and services on tank engine and transmission oil levels	3.297		
	CLUSTER CRITICALITY:		4.532	

CLUSTER 16: PLACE GUN TUBE IN TRAVEL LOCK

		M60A1	M48A5	M60A3
AA202	Place gun tube in travel lock	3.640	4.707	2.176
	CLUSTER CRITICALITY:		3.508	

CLUSTER 17: BORESIGHT OPTICS

		M60A1	M48A5	M60A3
AB217	Boresight Gunner's telescope	3.640	3.124	
AB219	Boresight IR sight of Gunner's periscope during daylight	4.528	2.972	
AB218	Boresight daylight sight of Gunner's periscope	3.851	2.972	
	CLUSTER CRITICALITY:		3.515	

CLUSTER 18: ASSEMBLE/INSTALL MACHINEGUNS*

TASK NO.		CRITICALITY M60A1	M48A5	M60A3
AB209	Clear an M219 machinegun	5.998	7.933	
AB210	Load an M219 machinegun	6.174	6.498	
AK208	Load a MAG-58 machinegun			5.345
AK207	Clear a MAG-58 machinegun			5.791
AB215	Mount an M219 machinegun in tank	6.361	6.098	
AK213	Mount a MAG-58 machinegun in tank			5.143
AA203	Remove the main gun breechblock group	5.726	5.921	5.558
AA208	Install the M37 periscope	4.002	5.146	4.039
AA205	Assemble the main gun breechblock	4.867	5.758	4.496
AB207	Assemble an M219 machinegun	5.265	5.291	
AK205	Assemble a MAG-58 machinegun			5.143
AB205	Clean an M219 machinegun	6.361	5.921	
AK203	Clean a MAG-58 machinegun			4.729

CLUSTER CRITICALITY: 5.565

CLUSTER 19: PERFORM OPERATIONAL CHECKS*

TASK NO.		M60A1	M48A5	M60A3
AB208	Check operation of an M219 machinegun	5.542	5.756	
AK206	Check operation of a MAG-58 machinegun			5.924
AA211	Determine corrective action required by replenisher tape	5.998	6.739	5.924
AA216	Perform emergency closing of main gun breech	6.377	7.026	6.071
AA209	Remove the M37 periscope	2.827	4.402	4.942
AA217	Perform before-operations maintenance checks on the tank generator blower	4.164	4.077	4.363
A3220	Perform after-operations maintenance checks and services on the air cleaners			3.473
A5203	Perform before-operations maintenance checks and services on air cleaners		3.900	
AD203	Perform before-operations maintenance checks and services on air cleaners	3.050		2.176

CLUSTER CRITICALITY: 4.881

GUNNER

TASK NO:	CLUSTER 1: ENGAGE TARGETS	CRITICALITY M60A1	M48A5	M60A3
A3302	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)			5.391
A3331	Zero tank main gun			6.792
AB304	Boresight IR sight of Gunner's periscope during daylight	4.735	5.244	
AB303	Boresight daylight sight of Gunner's periscope	5.000	5.244	
AL311	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary)(BEEHIVE)	5.776	5.878	
A3311	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary)(BEEHIVE)			5.615
A3305	Perform a zero pressure check (hydraulic power pack)			6.034
AA303	Boresight Gunner's telescope	6.075	5.392	5.237
AL315	Gunner fires main gun precision engagement using the TEL (stationary/moving)	6.322	6.322	
AL314	Gunner fires main gun precision engagement using the TEL (stationary/moving)	6.097	7.308	
AL303	Gunner fires main gun battlesight engagement using the GPD (stationary/moving)	5.523	6.522	
A3315	Gunner fires main gun precision engagement using the TEL (moving/moving)			5.615
A3303	Gunner fires main gun battlesight engagement using the GPD (moving/moving)			6.342
AK302	Zero mag-58 machinegun			4.970
A5305	Zero an M219 machinegun		4.600	
AF301	Zero an M219 machinegun	4.799		
AL317	Gunner fires main gun precision engagement using the GPD (stationary/stationary)(BEEHIVE)	5.922	5.392	
AL316	Gunner fires main gun precision engagement using the GPD (stationary/stationary)	6.097	6.272	
AL309	Gunner fires precision coax engagement using the TEL (stationary/stationary)	6.097	6.064	
AL305	Gunner fires nonprecision coax engagement using the TEL (stationary/stationary)	5.611	4.816	
AL302	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)	6.322	6.322	
A3317	Gunner fires main gun precision engagement using the GPD (moving/stationary)(BEEHIVE)			5.849
A3316	Gunner fires main gun precision engagement using the GPD (moving/stationary)			5.615
A3309	Gunner fires precision coax engagement using the TEL (moving/stationary)			4.593
A3305	Gunner fires nonprecision coax engagement using the TEL (moving/stationary)			6.034
A3313	Gunner fires coax rangecard lay to direct fire using the TEL (stationary/moving)			5.391
AL313	Gunner fires coax rangecard lay to direct fire using the TEL (stationary/moving)	4.587	5.244	
A3314	Gunner fires main gun precision engagement using the TEL (stationary/moving)			9.581
	CLUSTER CRITICALITY:		5.769	

	CLUSTER 2: PERFORM PREPARE-TO-FIRE PROCEDURES			
A1301	Perform main gun prepare-to-fire procedures from the Gunner's position	6.322		
A5302	Perform main gun prepare-to-fire procedures from the Gunner's position		6.842	
A0302	Perform main gun prepare-to-fire checks			6.582
	CLUSTER CRITICALITY:		6.582	

TASK NO:		M60A1	M48A5	M60A3
	CLUSTER 3: BORESIGHT SPECIAL SIGHTS			
A3326	Boresight M35E1 Gunner's periscope			4.760
AK306	Boresight TTS			5.391
	CLUSTER CRITICALITY:		5.076	
	CLUSTER 4: PREPARE RANGECARDS			
AA311	Prepare a sketch rangecard	5.052	4.371	3.893
AA312	Prepare a circular rangecard	4.476	4.524	4.549
	CLUSTER CRITICALITY:		4.478	
	CLUSTER 5A: OPERATE TURRET			
A3313	Operate gun elevating and turret traversing system in stabilized mode			5.391
A5301	Place turret into power operation		4.673	
A0301	Place turret into stabilized operation			4.593
AD301	Place turret into power operation	5.200		
	CLUSTER CRITICALITY:		4.837	
	CLUSTER 5B: PERFORM MISFIRE PROCEDURES			
AD303	Apply immediate action in case of main gun failure to fire	6.639		5.615
A5304	Apply immediate action in case of main gun failure to fire		6.842	
	CLUSTER CRITICALITY:		6.365	
	CLUSTER 6: ASSIST IN RANGECARD ENGAGEMENT			
A3310	TC fires main gun rangecard lay to direct fire using the RDS (stationary/stationary)(BEEHIVE)			4.549
AL312	TC fires coax rangecard lay to direct fire using the RFI (stationary/moving)	4.104	4.042	
AL310	TC fires main gun rangecard lay to direct fire using the RFD (stationary/stationary)(BEEHIVE)	4.626	4.042	
A3312	TC fires coax rangecard lay to direct fire using the RFI (stationary/moving)(BEEHIVE)			4.549
	CLUSTER CRITICALITY:		4.319	

TASK NO:		CRITICALITY M60A1	M48A5	M60A3
	CLUSTER 7: CONDUCT FIRE CONTROL INSTRUMENT CHECKOUT*			
AK303	Inspect tank thermal sight			5.391
A3319	Inspect Gunner's periscope M35E1			4.970
AA302	Prepare Gunner's telescope for operation	5.412	5.101	4.326
A3321	Perform XM21 computer self test procedures			5.391
A3327	Prepare tank for boresighting			5.615
A3325	Perform LRF malfunction detection test			5.849
AB306	Place ballistic computer into operation	5.643	5.704	
AK305	Perform TTS system test			5.391
A3323	Perform LRF self-test			5.849
AB302	Prepare Gunner's periscope for operation	5.746	5.101	
		CLUSTER CRITICALITY:	5.392	
	CLUSTER 8: BORESIGHT SEARCHLIGHT*			
AB307	Zero tank main gun	5.304	6.272	
A5307	Boresight tank searchlight using primary method		4.371	
AF303	Boresight tank searchlight using primary method	4.077		
AF302	Boresight tank searchlight using the alternate method (XENON)	2.049		
A5306	Boresight tank searchlight using the alternate method		4.371	
		CLUSTER CRITICALITY:	4.415	
	CLUSTER 9: ASSIST IN NIGHT .50 CALIBER ENGAGEMENT			
A3306	TC fires nonprecision .50 caliber engagement using the TPI (moving/moving)			4.593
AA310	Operate azimuth indicator	4.510	5.101	4.326
AL306	TC fires nonprecision .50 caliber engagement using the TPI (stationary/moving)	4.695	3.395	
		CLUSTER CRITICALITY:	4.437	
	CLUSTER 10: OPERATE ELEVATION AND GUNNER'S QUADRANT			
AA308	Operate Gunner's quadrant	4.695	4.214	3.149
AA307	Operate elevation quadrant	3.925	5.101	3.829
		CLUSTER CRITICALITY:	4.152	
	CLUSTER 11: PERFORM ZERO PRESSURE CHECK			
AA305	Perform a zero pressure check hydraulic power pack	5.412	5.101	4.970
		CLUSTER CRITICALITY:	5.161	

CLUSTER 12: PERFORM COMPUTER ELEVATION CHANNEL CHECK

TASK NO:		CRITICALITY M60A1	M48A5	M60A3
A3322	Perform XM21 computer elevation channel check			5.391

CLUSTER CRITICALITY: 5.391

CLUSTER 13: BORESIGHT MACHINEGUNS

TASK NO:		M60A1	M48A5	M60A3
AB308	Boresight an M219 machinegun mounted on a tank	4.253	3.853	
AK301	Boresight MAG-58 machinegun mounted on a tank			4.191

CLUSTER CRITICALITY: 4.099

CLUSTER 14: PREPARE AZIMUTH INDICATOR

TASK NO:		M60A1	M48A5	M60A3
AA309	Prepare azimuth indicator for operation	3.677	4.042	3.148

CLUSTER CRITICALITY: 3.622

CLUSTER 15: ASSIST IN TARGET ENGAGEMENTS*

TASK NO:		M60A1	M48A5	M60A3
A3329	Perform target range input (manual)			4.760
A3328	Perform target range input (laser)			5.079
AA301	Place turret into manual operation	5.000	4.673	2.733
AA313	Position gun tube in cradle in response to signals	2.049	3.853	3.148
AK307	Remove TTS			4.326
A3332	Remove M35E1 periscope image intensifier elbow, visible light elbow, and body assembly			4.760
AD303	Apply immediate action in case of main gun failure to fire	6.639		5.615
AA304	Set tank battlesights	4.799	4.887	4.549
A3330	Complete boresight procedures			3.829
AK304	Prepare tank thermal sight for operation			5.181
A3320	Prepare Gunner's periscope M35E1 for operation			5.391
A3301	TC fires main gun battlesight engagement using the RFD (moving/stationary)			5.422
AL308	TC fires precision coax engagement using the RFD (stationary/stationary)	4.253	3.645	
AL307	TC fires main gun precision engagement using the RFD (stationary/stationary)(BEEHIVE)	4.356	3.395	
AL304	TC fires nonprecision coax engagement using the RFI (stationary/moving)	4.476	3.395	
AL301	TC fires main gun battlesight engagement using the RFD (stationary/stationary)	4.695	4.673	
A3308	TC fires precision coax engagement using the RFD (stationary/stationary)			5.391
A3307	TC fires main gun precision engagement using the RFD (moving/stationary)(BEEHIVE)			5.391
A3304	TC fires nonprecision coax engagement using the RFI (moving/moving)			5.181

CLUSTER CRITICALITY: 4.481

CLUSTER 16: DRAIN REPLENISHER SYSTEM

TASK NO:		M60A1	M48A5	M60A3
AA306	Drain replenisher system	4.104	4.042	4.515

CLUSTER CRITICALITY: 4.220

TASK NO:		CRITICALITY M60A1	CRITICALITY M48A5	CRITICALITY M60A3
	CLUSTER 17: INSTALL/TEST SIGHTING SYSTEMS*			
AK308	Install TTS			4.191
A3333	Install M35E1 periscope image intensifier elbow, visible light elbow, and body assembly			4.970
AK309	Activate muzzle reference system			4.549
	CLUSTER CRITICALITY:		4.570	
	CLUSTER 18: PREPARE TANK FOR BORESIGHTING			
AB301	Prepare tank for boresighting	5.264	5.244	
	CLUSTER CRITICALITY:		5.254	
	CLUSTER 19: FILL REPLENISHER			
A5303	Fill replenisher system		4.673	
AD302	Fill replenisher system	4.587		3/776
	CLUSTER CRITICALITY:		4.345	
	CLUSTER 20: PERFORM CHECKS AND SERVICES ON PERISCOPE			
A3334	Perform before-operations maintenance checks and services on periscope M35E1			5.391
A3335	Perform after-operations maintenance checks and services on periscope M35E1			5.391
	CLUSTER CRITICALITY:		5.391	

TANK COMMANDER

TASK NO:	CLUSTER 1: OPERATE WEAPON SYSTEMS	CRITICALITY M60A1	M48A5	M60A3
A3419	Inspect Tank Commander's periscope M36E1			4.437
AD402	Load an M85 machinegun mounted on a tank	5.361		4.588
AK403	Perform TTS system test			4.437
AK401	Inspect tank thermal sight			4.875
A3421	Perform LRF self-test			4.282
A5404	Clear an M2 machinegun		3.781	
A5414	Apply immediate action to reduce stoppage of an M2 machinegun		4.912	
AD408	Apply immediate action to reduce stoppage of an M85 machinegun	5.203		6.102
A5412	Unload an M2 machinegun mounted on a tank		4.176	
AD404	Unload an M85 machinegun	3.684		5.155
AD403	Clear an M85 machinegun	5.046		5.916
A5402	Mount an M2 machinegun in a tank		4.176	
AD410	Remove an M85 machinegun from a tank	4.731		3.723
AD401	Mount an M85 machinegun in a tank	4.193		4.875
AD407	Assemble an M85 machinegun	5.528		4.776
AB402	Boresight rangefinder with the main gun bore axix alined on an aiming point at 1200 meters	5.703	6.709	
AA401	Prepare tank for boresighting	5.203	5.076	5.916
A5407	Zero an M2 machinegun		5.150	
AF402	Zero an M85 machinegun	5.361		
A3423	Boresight LRF sight			5.442

CLUSTER CRITICALITY: 4.951

	CLUSTER 2: ADJUST HEADSPACE AND TIMING	M60A1	M48A5	M60A3
A5409	Adjust headspace on the M2 machinegun		5.150	
A5410	Adjust timing on the M2 machinegun		5.638	

CLUSTER CRITICALITY: 5.394

	CLUSTER 3: INSTALL AND REMOVE EQUIPMENT	M60A1	M48A5	M60A3
AB405	Disassemble an M85 machinegun	4.193		5.015
AK404	Remove TTS			3.477
A3428	Remove 36E1 periscope image intensifier elbow visible light elbow, and body assembly.			3.477
A5408	Remove an M85 machinegun from a tank		4.349	
A5405	Disassemble an M85 machinegun		4.994	
A3429	Install M36E1 periscope image intensifier elbow visible light elbow, and body assembly			3.477
A3430	Remove periscope M36E1 head assembly			2.704
A3422	Install safety filter on receiver/transmitter unit			2.704
A3425	Perform target range input (manual)			5.296
AK405	Install TTS			4.437

CLUSTER CRITICALITY: 4.011

	CLUSTER 4: PERFORM TARGET RANGE INPUT (LASER)	M60A1	M48A5	M60A3
A3424	Perform target range input (laser)			5.296

CLUSTER CRITICALITY: 5.296

TASK NO:	CLUSTER 5: PERFORM MAIN GUN PREPARE-TO-FIRE PROCEDURES	CRITICALITY M60A1	M48A5	M60A3
AF401	Perform main gun prepare-to-fire procedure from the TC's position	5.703		
A5401	Perform main gun prepare-to-fire procedure from the TC's position		5.150	
	CLUSTER CRITICALITY:		5.427	

TASK NO:	CLUSTER 6: PERFORM TACTICAL GUNNERY PROCEDURES	M60A1	M48A5	M60A3
A3406	TC fires nonprecision .50 caliber engagement using the TPI (moving/moving)			5.593
AL404	TC fires nonprecision coax engagement using the RFI (stationary/moving)	4.193	3.988	
A3404	Zero tank main gun			5.916
AL408	TC fires precision coax engagement using the RFD (stationary/stationary)	4.193	4.176	
AL406	TC fires nonprecision .50 caliber engagement using the TPI (stationary/moving)	4.193	5.307	
A3408	TC fires precision coax engagement using the RFD (stationary/stationary)			6.553
A3401	TC fires main gun battlesight engagement using the RFD (moving/stationary)			6.553
AL407	TC fires main gun precision engagement using the RFD (stationary/stationary)(BEEHIVE)	6.408	7.060	
A3407	TC fires main gun precision engagement using the RFD (moving/stationary)(BEEHIVE)			6.102
AL401	TC fires main gun battlesight engagement using the RFD (stationary/stationary)	5.899	6.709	
A3414	Gunner fires main gun precision engagement using the TEL (stationary/moving)			6.553
AL410	TC fires main gun range card lay to direct fire using the RFD (stationary/stationary)(BEEHIVE)	5.046	4.837	
A3410	TC fires main gun rangecard lay to direct fire using the RFD (stationary/stationary)(BEEHIVE)			5.916
A3402	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)			5.442
AL417	Gunner fires main gun precision engagement using the GPD (stationary/stationary)(BEEHIVE)	5.528	6.435	
AL416	Gunner fires main gun precision engagement using the GPD (stationary/stationary)	5.361	5.999	
AL415	Gunner fires main gun precision engagement using the TEL (stationary/moving)	5.985	5.150	
AL414	Gunner fires main gun precision engagement using the TEL (stationary/moving)	6.408	4.761	
AL409	Gunner fires precision coax engagement using the TEL (stationary/stationary)	4.107	4.681	
AL405	Gunner fires nonprecision coax engagement using the TEL (stationary/stationary)	3.374	4.030	
AL403	Gunner fires main gun battlesight engagement using the GPD (stationary/moving)	6.509	5.182	
AL402	Gunner fires main gun battlesight engagement using the GPD (stationary/stationary)	5.899	5.999	
A3417	Gunner fires main gun precision engagement using the GPD (moving/stationary)(BEEHIVE)			5.748
A3416	Gunner fires main gun precision engagement using the GPD (moving/stationary)			6.307
A3415	Gunner fires main gun precision engagement using the TEL (moving/moving)			5.916
A3409	Gunner fires precision coax engagement using the TEL (moving/stationary)			5.162
A3405	Gunner fires nonprecision coax engagement using the TEL (moving/stationary)			4.875
A3403	Gunner fires main gun battlesight engagement using the GPD (moving/moving)			6.102
	CLUSTER CRITICALITY:		5.480	

TASK NO:	CLUSTER 7: TROUBLESHOOT MACHINEGUNS	M60A1	M48A5	M60A3
AD409	Troubleshoot an M85 machinegun using TM 9-2350-215-10, Table 3-6	4.979		4.437
A5415	Troubleshoot an M2 machinegun using TM 9-2350-258-10, Table 3-6		2.928	
	CLUSTER CRITICALITY:		4.115	

TASK NO:	CLUSTER 8: TROUBLESHOOT MACHINEGUNS	M60A1	M48A5	M60A3
A5406	Assemble an M2 machinegun		3.781	
	CLUSTER CRITICALITY:		3.781	

TASK NO:	CLUSTER 9A: BORESIGHT AND ZERO WEAPONS	CRITICALITY M60A1	CRITICALITY M48A5	CRITICALITY M60A3
AB406	Zero tank main gun	5.985	7.021	
A3434	Zero tank main gun			6.867
AB406	Boresight tank searchlight using the alternate method (XENON)	1.599	4.349	
AB403	Determine range to target with rangefinder (coincidence)	6.408	5.811	
A3427	Zero M85 machinegun			5.155
AB401	Prepare tank rangefinder for operation	6.408	7.572	
A3426	Boresight M85 machinegun mounted on a tank			5.593
A5411	Boresight M2 machinegun mounted on a tank		5.470	
AF403	Boresight M85 machinegun mounted on a tank	4.731		
		CLUSTER CRITICALITY:	5.613	

	CLUSTER 9B: FIRE RANGECARD ENGAGEMENT			
A3412	TC fires coax rangecard lay to direct fire using the RFI (stationary/moving)			5.296
AL412	TC fires coax rangecard lay to direct fire using the RFI (stationary/moving)	5.361	5.150	
		CLUSTER CRITICALITY:	5.269	

	CLUSTER 10: OPERATE TANK RADIO			
AA406	Operate tank radio	4.846	4.349	5.015
AA407	Perform operational checks on tank radios	4.889	4.518	4.677
		CLUSTER CRITICALITY:	4.716	

	CLUSTER 11: ASSIST IN RANGECARD ENGAGEMENTS			
A3411	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary)(BEEHIVE)			5.593
AL413	Gunner fires coax rangecard lay to direct fire using the TEL (stationary/moving)	4.193	3.738	
AL411	Gunner fires main gun rangecard lay to direct fire using the GPD (stationary/stationary)(BEEHIVE)	5.046	5.470	
A3413	Gunner fires coax rangecard lay to direct fire using the TEL (stationary/moving)			4.875
		CLUSTER CRITICALITY:	4.819	

	CLUSTER 12: ILLUMINATE TARGETS			
AB408	Illuminate targets using tank searchlight	3.968	4.518	
		CLUSTER CRITICALITY:	4.243	

	CLUSTER 13: PREPARE RANGECARDS			
AA403	Prepare a sketch rangecard	5.046	4.837	5.748
AA404	Prepare a circular rangecard	3.684	3.988	3.928
		CLUSTER CRITICALITY:	4.539	

TASK NO:		CRITICALITY M60A1	CRITICALITY M48A5	CRITICALITY M60A3
	CLUSTER 14: BORESIGHT SEARCHLIGHT			
AB407	Boresight tank searchlight using primary method	3.968	3.781	
	CLUSTER CRITICALITY:		3.875	
	CLUSTER 15: ACQUIRE TARGETS			
AA402	Acquire ground targets (night)	6.819	5.470	6.867
	CLUSTER CRITICALITY:		6.385	
	CLUSTER 16: OPERATE SEARCHLIGHT			
AB405	Place tank searchlight into operation	3.968	4.212	
	CLUSTER CRITICALITY:		4.090	
	CLUSTER 17: PREPARE OPTICAL EQUIPMENT FOR OPERATION			
AK402	Prepare tank thermal sight for operation			3.723
A3420	Prepare TC's periscope M36E1 for operation			4.114
	CLUSTER CRITICALITY:		3.919	
	CLUSTER 18: ACTIVATE SMOKE GRENADE LAUNCHER			
AK406	Activate smoke grenade launcher			4.734
	CLUSTER CRITICALITY:		4.734	
	CLUSTER 19: INSTALL AND MAINTAIN OPTICAL EQUIPMENT*			
A3431	Install periscope M36E1 head assembly			3.723
A3418	Operate gun elevating and turret traversing system in stabilized mode			4.114
A3432	Conduct before-operations maintenance checks and services on periscope M36E1			3.723
A3433	Conduct after-operations maintenance checks and services on periscope M36E1			3.928
AA405	Set tank battlesights	5.046	5.470	5.155
	CLUSTER CRITICALITY:		4.451	
	CLUSTER 20: SERVICE MACHINEGUNS			
AD406	Service an M85 machinegun	5.046		4.588
A5413	Service an M2 machinegun mounted on a tank		3.988	
	CLUSTER CRITICALITY:		4.541	

APPENDIX C

EXPLANATION OF TASK CODE NUMBERS

EXPLANATION OF TASK CODE NUMBERS

Each task was identified by a five-place alpha-numeric code. The first two of the five places identify tasks whose performance is common or unique to the tanks, as shown in the following table:

	TANK SYSTEMS			
Designators	M60A1	M60A1(AOS)[1]	M48A5	M60A3
AA	X	X	X	X
AB	X	X	X	
AD	X	X		X
AF	X	X		
AL	X		X	
AO		X		X
A1	X			
AS		X		
A3				X
A5			X	
AK				X(NEW)

Task numbers beginning with AA indioate tasks whose performance is common to all four tanks; those beginning with A1 are unique to the M60A1, and so forth.

The third place in the code is a numeral indicating duty positions as follow: 1 = Driver, 2 = Loader, 3 = Gunner, 4 = Tank Commander.

The numbers in the last two places simply distinguish tasks within the various tank/duty position categories; A5103, for example, is task number 3 in the M48A5 Driver set.

[1]Task lists for the M60A1(AOS), though not contractually required, were prepared because doing so required little effort. They were not used in subsequent analyses.

APPENDIX D

METHOD FOR PAIRING TASKS IN THE PARTIAL PAIRED COMPARISON QUESTIONNAIRES

METHOD FOR PAIRING TASKS IN THE PARTIAL PAIRED COMPARISON QUESTIONNAIRES

The method followed for pairing tasks had three steps:

(1) Decide how many times to pair each task. This decision is governed by the amount of time respondents can devote to the study. The rule for this study was: If the task list has an even number of tasks, pair each task seven times; if the task list has an odd number of tasks, pair each task six times.

(2) Calculate the total number of pairs desired. The formula for this calculation is:

$$\frac{\text{Tasks on list} \times \text{Number of pairs}}{2} = \text{Total pairs desired.}$$

(3) Select random tasks for pairing. This step requires a four part procedure:

- Determine an interval by dividing the number of tasks by the desired number of pairs.
- Select the first starting point (or points) for counting. If the number of tasks is even, start at the approximate midpoint of the task list. If the number of tasks is odd, start at the two points that bracket the midpoint by half the interval.
- Count out from the starting point (or points) and select the starting point and each task at the interval to be paired with Task 1.
- To select pairs for succeeding tasks add one to each task number paired with the preceding task.

Stop pairing tasks when the desired number of pairs is reached.

This method of forming the pairs may be illustrated by two examples. The total number of tasks for the M60A1 Driver was 70. Since the total number of tasks is even, seven pairings of each are desired. The total number of pairs of tasks that will appear on the questionnaire is $\frac{70 \times 7}{2} = 245$.

An interval is obtained by dividing the number of tasks by the desired number of pairings for each task: $70 \div 7 = 10$. One then begins at the approximate midpoint of the 70 tasks, using the interval to count up and down from the midpoint to obtain seven task numbers. The seven task numbers thus obtained are 35 (approximate midpoint), 25 (ten less than 35), 15 (another ten less), 5 (another ten less); 45 (ten more), 55, and 65. The tasks corresponding to these numbers are paired with Task 1. Task 2 is paired with the seven tasks corresponding to each of the seven task numbers plus one: Task 2 is paired with Task 6, then with 16, with 26, and so forth. Task 3 is paired with each of the seven numbers for Task 2 plus one: 3 with 7, 3 with 17, 3 with 27, and so forth. The progression is followed until the desired number of pairs (245 in this case) is reached.

If the total number of tasks is odd and six pairings of each are desired, a procedure is followed that is identical in most respects to the one described above. The difference is that after obtaining the interval, one begins counting up and down, not from the approximate midpoint, but from two points approximately equidistant by half the interval from the midpoint. For example, the total number of tasks for the M60A3 Loader was 65. The number of pairs of tasks that will appear on the questionnaire is $\frac{65 \times 6}{2} = 195$. The interval is $65/6 = 11$, and the midpoint is 33. Adding and subtracting approximately half the interval to and from the midpoint yield starting points at Tasks 27 and 38 (or 28 and 39). Counting up and down by 11 yields four additional tasks (numbers 5, 16, 49, and 60). These and Tasks 27 and 38 get paired with Task 1. Task 2 is paired with Tasks 6, 17, 28, 39, 50, and 61; and so forth until the desired number of pairs (195) is reached.

The methods described above are applicable in all cases where the total number of tasks is greater than 28. At some numbers of tasks less than 28, the effects of rounding the interval present problems. With a total of 20 tasks, for example, Task 1 would get paired with itself. And with a total of 10 tasks, the interval is one, which would lead to a complete rather than a partial pairing of tasks. These problems are unimportant, since with a small number of tasks, the use of complete pairings would become feasible and the need for using partial pairings would disappear.

APPENDIX E

INSTRUCTIONS TO RESPONDENTS
FOR THE PAIRED COMPARISON
QUESTIONNAIRES

INSTRUCTIONS TO RESPONDENTS FOR THE
PAIRED COMPARISON QUESTIONNAIRE

Materials

Please check to see that you have two sets of papers in addition to these instructions. The two sets of papers are:

A. A set of Answer Sheets,* and

B. A set of papers entitled "Paired Compairsons."

If you do not have both sets of papers, please raise your hand and we'll give you what you need.

Personal Data

Please look at the cover page of the Answer Sheets, entitled "Personal Data." We'd like you to fill in your name, rank, and so forth. Please be assured that your answers will be treated as anonymous. Our interest is not in who gives what answers, and none of this information will be used against you. Later on though, we may want to find out if people with different kinds and amounts of experience answered the questions differently. We also may want to contact you for some follow-up questions. To do these things we will need the Personal Data.

Please fill in all the blanks on the cover page of the Answer Sheets. If anything is not clear, please ask questions.

Purpose of the Exercise

The purpose of this exercise is to find out what sorts of priorities you place on crew members' ability to perform various tasks. To do this, we would like you to make several assumptions:

---

*Last-minute changes required not using answer sheets, and that the questionnaires be taken home by respondents rather than administered in a conference room as originally intended. Respondents were told, therefore, to circle their responses on the questionnaire, and to ignore parts of the instructions that implied group administration.

- Assume that you are a company commander.
- Assume further that you must choose crew members to take on a mission.
- Assume also that you and your crews are certain to encounter the enemy during the mission, and will exchange fire with him.

To get you to choose crew members, we will present several pairs of tasks. The crew member whom you choose can do only one of the two tasks in each pair. Each of you will be dealing with only one crew position and only one tank. Here's an example of a pair of tasks like the ones we'll ask you about:

A. Inspect an M219 machinegun.

B. Stow main gun rounds in tank.

(The example is for an M60A1 Loader, which may not correspond to the tank and crew position that you'll be dealing with. But the instructions that follow apply regardless of the tank and crew position that you'll be working with.)

If you choose A in the example, you will get a Loader who can inspect an M219 machinegun, but cannot stow main gun rounds in an M60A1. If you choose B in the example, you will get a Loader who can stow main gun rounds in the M60A1, but cannot inspect an M219 machinegun. (We realize that this is not a realistic assumption, but please accept it for purposes of the study.)

Any questions up to this point? If so, raise them now, and let's try to get them answered. If not, please proceed with the following five practice problems. All of the practice problems apply to the M60A1 Loader. The problems that you will do later may apply to a different tank and a different crew position.

Practice Problems

P1

A. Mount an M219 machinegun in tank.

B. Perform operator maintenance on radios and accessories.

If you would rather have the Loader who can mount an M219 machinegun, darken A in the P1 row of the Practice block of the Answer Sheet. If you would rather have the Loader who can perform operator maintenance on radios and accessories, darken B in the P1 row. Please make your marks dark and heavy. The answer sheets will be machine scored.

P2

A. Clean an M219 machinegun.

B. Boresight IR sight of Gunner's periscope during daylight.

Would you rather have a Loader who could do A, or a Loader who could do B? Remember -- you can't have both, so you must choose one. If A, darken A after P2 on the Answer Sheet. If B, darken B. Any questions up to this point? If so, please raise them. If not, please complete practice problems P3, P4, and P5:

P3

A. Install main gun breechblock.

B. Service tank main gun ammunition.

P4

A. Unload misfired main gun round.

B. Disassemble the breechblock.

P5

A. Operate vehicular intercommunications equipment.

B. Place gun tube in travel lock.

If you've completed all five practice problems and have no questions, please read the section that follows, and then proceed with the remaining items. Take your time, and if there's any part of the exercise you don't understand, please ask us about it.

Note on Gunnery Items

Several of the comparisons that you will make will involve gunnery items, which require a word of explanation. Here's a pair of gunnery tasks for the M60A1:

A. Gunner fires main gun battlesight engagement using the GPD (stationary/moving).

B. Tank Commander fires nonprecision .50 caliber engagement using the TPI (stationary/moving).

The fire control instruments in this example and in all the other gunnery items will be abbreviated. The abbreviations and their definitions are:

AUX = Auxiliary Fire Controls

GPD = Gunner's Periscope Day

GPI = Gunner's Periscope Infrared

INF = Infinity Sight

RFD = Rangefinder Day

RFI = Rangefinder Infrared

TEL = Telescope

TPD = Tank Commander's Periscope Day

TPI = Tank Commander's Periscope Infrared

The two words in parentheses after each item refer to the movement of the firing vehicle and the target -- in that order. Thus, moving/stationary means moving firing vehicle/stationary target. And stationary/moving means stationary firing vehicle/moving target.

Finally, all gunnery items begin with either the word Gunner or Tank Commander. This does not necessarily mean that you are choosing a Gunner or a Tank Commander. Suppose, for example, that the notation at the top of your paired comparison sheet is for Loader, M60A1. And you have a gunnery item such as:

A. Gunner fires main gun battlesight engagement using the GPD (stationary/moving).

B. Tank Commander fires nonprecision .50 caliber engagement using the TPI (stationary/moving).

If your job is to choose a Loader, you must ask yourself, "Would I rather have a <u>Loader</u> who could perform the <u>Loader's</u> duties associated with A above; or a <u>Loader</u> who could perform the <u>Loader's</u> duties associated with B, above?" The fact that the Gunner is firing one of the engagements in the example, and the Tank Commander is firing the other engagement is largely irrelevant here, since we're choosing not a Gunner or a Tank Commander, but a Loader.

APPENDIX F

PLAN FOR EXAMINING CONSTRUCT VALIDITY OF THE CRITICALITY RATINGS

PLAN FOR EXAMINING CONSTRUCT VALIDITY OF THE CRITICALITY RATINGS

The main requirement in any plan to validate skill criticality ratings is to minimize dependence on expert judgment in defining the criterion measures. If this is not done, then validation reduces to establishing the correlation between two sets of expert opinions. High correlations might indicate reliable ratings (that both sets of ratings were made on the same or highly correlated concepts), but are not adequate evidence that judges were considering the concept of criticality in their ratings.

The ideal validation plan would involve actual or simulated combat missions, embarked upon under identical conditions as many times as there are identified skills. On each enactment, one skill would be missing. Attainment of the mission objective would then be rated as success or failure. By replicating across many missions, the proportion of failures would be used as the criticality rating for the skill designated as "missing" for those mission enactments.

Such an approach would certainly provide information concerning the degree to which deficiencies in skills degrade performance of a mission, or criticality. But the disadvantages are obvious and overwhelming: time and cost requirements; impossibility of standardizing conditions; and difficulty in ensuring that tasks in all skill areas are performed adequately, except for those in the "missing" skill, which must not be performed. If the tasks and skills could be fully defined in terms of initiators, standards of performance, and consequences of performance or nonperformance, and if all interactions among consequences of performance or nonperformance of all skills were known, and if all consequences and interactions of consequences could be empirically related to success or failure, then a mathematical model could be defined and computer-simulated to overcome all the former difficulties. This would be a major task, for which data concerning "successful" consequences would have to be obtained as described above, at which point the same disadvantages immediately would re-emerge. The need for actual or simulated missions could be side-stepped by presenting the situations to a panel of experts and obtaining their judgments of specific consequences of inadequate performance on each skill, which could then be converted to, perhaps, a five-point success/failure scale. This again reduces to a set of expert opinions, which may reflect task difficulty or frequency of performance as well as criticality.

From the foregoing it may be seen that there are two general approaches to obtaining skill criticality ratings for purposes of validation: the empirical study, to obtain "real" criticality, or the expert questionnaire study, to obtain estimates of criticality. The first is costly, time-consuming, and practically (as opposed to theoretically) impossible. The second produces results which, though possibly reliable, may be confounded among criticality, difficulty, complexity, or frequency of performance. Any combination of the two approaches, while it may serve to eliminate some of the problems inherent in one, will necessarily be subject to problems associated with the other.

A method is available, however, whereby the expert ratings of criticality, obtained through the paired-comparison technique, may be examined for possible influences or contamination from factors other than criticality. The correlational study of validity, developed by Campbell and Fiske (1959), encompasses measures of several factors, each measured by two or more methods. Measures of the same factor by dissimilar methods should converge, while measures of different factors by the same or different methods should diverge.

The most frequently encountered challenges to the validity of criticality ratings are that the ratings represent learning difficulty (DF), or performance deficiency (PD) as perceived by raters. The validation study will examine skill ratings as derived from task ratings on these variables and on criticality (CR) by two methods. The results of the analysis will provide information concerning the independence of the criticality variable from other variables that might influence criticality ratings.

METHOD

Raters

The measures of criticality and other variables will be obtained from volunteers from the Armor Officers' Advanced Course at Fort Knox. Each person will respond to items by the two methods for criticality, difficulty to learn, and performance deficiency.

Procedure 1: *Paired Comparisons*

The first method will require raters to make judgments of the criticality (CR), learning difficulty (DF), and performance deficiency (PD) of pairs of tasks. Twenty tasks will be paired according to the partial-pairing algorithm of McCormick and Bachus (1952), yielding a total of 60 pairs to be judged three times in each of the twelve sets. On the basis of the raters' judgments, scale values for CR, DF, and PD will be assigned to each of the tasks judged. These values will then be averaged for tasks within the skill clusters defined by the cluster analysis, across tanks, to yield CR, DF and PD scale values for each skill within the four duty positions, for each rater.

Tasks

Each of the twelve sets of tasks will be comprised of a sample of all tasks from each duty position (Driver, Loader, Gunner, Tank Commander) by each tank (M60A1, M48A5, M60A3). The tasks were assigned criticality ratings in the paired comparison study described in this report. A total of 20 tasks from the criticality study will be used in the validation. The 20 tasks will be the seven most critical, the seven least critical, and the six closest to the median criticality rating.

Instructions

To obtain the CR ratings, the same instructions will be given to the raters as were given in the criticality study.

In obtaining ratings of DF, the instructions to the raters will vary only in that they are instructed to assume that they must decide which of the two crew members, each of whom is deficient on one task, will require the greatest amount of practice in order to bring him up to proficiency on that task, so that he would be able to perform the task adequately in a live fire engagement.

For ratings of PD, the instructions will ask the raters to judge on which of a pair of tasks incumbents are more likely to be deficient.

By this method, each of three factors -- CR, DF, and PD -- has an implicit operational definition, as follows:

CR (criticality) - the extent to which deficiency on the task would degrade mission success.

DF (learning difficulty) - the amount of practice needed to ensure proficiency on a task.

PD (performance deficiency) - likelihood that incumbents are deficient on the task.

Each of the raters will make judgments for all three dimensions, on only one of the 12 sets of tasks (four duty positions within each of three tanks). At least five raters must rate each of the sets.

Procedure 2: Rating of Behavioral Descriptors

Each task considered in this study already has been characterized in terms of a set of task descriptors. These descriptors will be rated by the raters in terms of CR, DF and PD. The ratings will then be summed for each task, according to whether or not the descriptor is involved in performance of the task, and then averaged for tasks within the skill clusters to yield scale values for CR, DF and PD within each duty position for each rater.

Behavioral Descriptors

The behavioral descriptors to be used in the ratings are those that were used to define the tasks for the cluster analyses.* They are listed and defined in Appendix A.

Instructions

The raters will be given the list of behavioral descriptors and a list of the definitions of the descriptors. They will be instructed to rate the 32 tasks on a scale from 1 to 50, on CR, DF, and PD, where 1 = least critical/difficult/deficient, and 50 = extremely critical/difficult/deficient. The three factors will be defined for the raters as:

CR - the extent to which deficient performance on the descriptor would degrade performance of the soldier's tasks.

DF - the amount of practice required by the soldier to attain proficiency on the behavior.

PD - the likelihood that incumbents will be deficient in performance of the behavior.

---

*Only 32 of the descriptors will be used. The descriptors numbered 8 (Smell), 17 (None), 24 (Identifies Symbols) and 36 (None) will be deleted because they were not used to characterize any task in the original study.

The instructions will be similar to those shown in Appendix I. Each rater will consider the descriptors relative to only one of the four duty positions, the same duty position which he considered in making the paired comparison ratings. Thus the descriptors will be considered by at least 15 raters for each duty position.

ANALYSIS

The first step in the analysis will be to compute a rank order correlation between the CR values obtained from the paired comparisons in the Criticality Study and in the Validation Study. All skills will be ranked from 1 to N (the number of skills for the duty position) on the two sets of CR values; the rank order correlation should be at least .60 to ensure that the same construct of criticality is being validated.

For each of the four sets of skills (one for each duty position), the scale values of CR, DF, and PD from each rater by the two methods will be correlated. The correlations will be entered in a correlation matrix, as illustrated in Table H-1.

The hypothesis is that the correlations will be fairly substantial in the sections of the matrix for each variable by the two methods (superscribed a, b, and c in Table H-1, and that the remaining correlations, which presumably pair distinctive variables, will be low. The measures of CR and PD converge very well in the example, having correlations of .91 and .89, respectively. The two measures of DF correlate somewhat lower (.75), but still higher than ratings of different variables by the same methods (superscribed d and e). The correlations between DF and CR by either method are only slightly higher than within-method correlations between DF and PD but considerably higher than the within-method correlations between CR and PD. This suggests that DF is more difficult for raters to assess than CR or PD, and somewhat more easily confused with CR than

TABLE F-1

MULTIFACTOR-MULTIMETHOD MATRIX OF HYPOTHETICAL CORRELATIONS OF CRITICALITY, LEARNING DIFFICULTY, AND PERFORMANCE DEFICIENCY SCALE VALUES OBTAINED BY PAIRED COMPARISONS AND RATINGS OF BEHAVIORAL DESCRIPTORS

FACTOR		CR		DF		PD	
	METHOD	1	2	1	2	1	2
CR	1	–	.91[a]	.31[e]	.16	.26[e]	.10
	2		–	.18	.32[d]	.12	.29[d]
DF	1			–	.75[b]	.30[e]	.21
	2				–	.19	.31[d]
PD	1					–	.89[c]
	2						–

is PD. Still, each of the three variables emerges as distinct, with little overlap between variables within methods, and high convergence within variables across methods.

The data obtained in the administration of the two instruments for each of the three variables will be entered into multivariable-multimethod matrices for each set of skills. The matrices will then be examined for convergence and divergence as described and illustrated in the example.

The validity of the criticality ratings can, of course, be challenged on the grounds of confounding by sources other than learning difficulty and performance deficiency. The effects of the other sources can be isolated using a design identical to the one described here.

APPENDIX G

DEFINITIONS OF TASK DESCRIPTORS

DEFINITIONS OF TASK DESCRIPTORS

STIMULI

1. Written (textual) material: (books, job instructions, signs, technical manuals.)
2. Graphic/tabular material: (Materials which deal with quantities or amounts and displayed in graphic or tabular form.)
3. Instrument read-outs: (Tools, equipment, machinery which are sources of information when observed during use or operation, for example, dials, gauges, signal lights, radarscopes, speedometers, timing light, mine detector, multimeter.)
4. Natural environmental features: (Landscapes, fields, geological samples, vegetation, cloud formations, and other features of nature which are observed or inspected to provide information.)
5. Man-made environmental features: (Man-made or altered aspects of the indoor or outdoor environment which are observed or inspected to provide job information; do not consider equipment or machines that a soldier uses in his work. For example, structures, buildings, dams, highways, bridges, docks, railroads.)
6. Oral command or request: (Verbal orders, instructions, requests, conversations, interviews, discussions, formal meetings. Consider only verbal communication that is relevant to performance.)
7. Non-verbal sounds; (Noises, engine sounds, sonar, signals, horns.)
8. Smell (olfaction): (Odors which the soldier needs to smell in order to initiate performance; do not include odors simply because they happen to exist in the work environment.)
9. Body feel (kinesthesis): (Sensing or recognizing changes in the direction or speed at which the body is moving without being able to sense them by sight or hearing.)

10. Touch: (Pressure, pain, temperature, moisture; provides information stimulus for performing the task.)

11. Self-initiated: (If a task can be performed without performing a sub-task, no matter the consequences of not performing the sub-task, then that sub-task is self-initiated. For example, the Loader can LOAD TANK MAIN GUN without "checking replenisher tape," "inspecting the chamber for obstruction," or "standing clear of path of recoil." These sub-tasks are then self-initiated.)

TOOLS, INSTRUMENTS, AND CONTROLS

12. Common hand tools and measuring devices: (Tools used to perform operations not requiring great accuracy or precision; for example, hammers, wrenches, trowels, knives, scissors, chisels, putty knives, strainers, hand grease guns. Measuring devices include rules, measuring tapes, micrometers, calipers, protractors, squares, thickness gauges, levels, volume measuring devices, tire gauges. Tools and measuring devices which are not unique to a tank environment.)

13. Special hand tools and measuring devices: (Tools and measuring devices which are unique to a tank environment. For example, the extracting and ramming device.)

14. Activation controls: (Hand-or foot-operated devices used to start, stop, or otherwise activate energy-using systems or mechanisms. For example, light switches, electric motor switches; ignition switches, power turret traverse.)

15. Fixed setting controls: (Hand- or foot-operated devices with distinct positions, detents, or definite settings. For example, gearshift, machinegun safety switch, ammunition control handle.)

16. Variable setting controls: (hand- or foot-operated devices that

can be set at the beginning of operation, or infrequently, at any position along a scale. For example, TV volume control, room thermostat, rheostat, rangefinder range knob.)

17. None: (Tools, instruments, or controls are not used when performing the task on sub-task.)

MEDIATING PROCESSES

18. Recalls bodies of knowledge: (Concerns verbal or symbolic learning; acquisition and long-term maintenance of knowledge so that it can be recalled. For example, recalling equipment nomenclature or functions, recalling system functions, recalling specific radio frequencies and other discrete facts.)

19. Uses verbal information: (Concerns the practical application of information, limited uncertainty of outcome, little thought of other alternatives. For example, based on academic knowledge: determine which equipment to use for a specific task; compare alternative modes of operation of a piece of equipment and determine the appropriate mode for a specific situation. Based on memorized knowledge of radio frequencies, choose the correct frequency in a specific situation.)

20. Uses rules: (Choosing a course of action based on applying known rules, frequently involves "if ... then" situations. The rules are not questioned, the decision focuses on whether the correct rule is being applied. For example, apply the "rules of the road," solve mathematical equations, select proper fire extinguisher for different type fires.)

21. Makes decisions: (Choosing a course of action when alternatives are unspecified or unknown; a successful course of action is not readily apparent. The penalties for unsuccessful courses of

action are not readily apparent. Frequently involves forced decisions made in a short period of time with soft information. For example, threat evaluation and weapon assignment; choosing a diagnostic strategy in dealing with a malfunction in a complex piece of equipment.)

22. Detects (including vigilance): (Vigilance -- detect a few cues embedded in a large block of time. Low threshold cues; early awareness of small cues. For example, early detection of a target, detect, through a slight change in sound, a bearing starting to burn out in a power generator.)

23. Classifies: (Pattern recognition approach of identification -- not problem solving. Classification by non-verbal characteristics. Object to be classified can be viewed from many perspectives or in many forms. For example, classify a target as "friendly" or "enemy"; determine that an identified noise is a wheel bearing failure, not a water pump failure by rating the quality of the noise -- not by the problem solving approach.)

24. Identifies Symbols: (Involves the recognition of symbols which typically are of low meaningfulness to untrained persons. Identification, not interpretation, is emphasized. Involves storing queries of symbolic information and related meanings. For example, reading electronic symbols on a schematic drawing; identifying map symbols; reading and transcribing symbols on a tactical status board.)

25. Recalls set procedures: (Concerns the chaining or sequencing of events; includes both the cognitive and motor aspects of equipment set-up and operating procedures. Need to follow specific set procedures on routines in order to obtain satisfactory outcomes. For example, recalling equipment assembly and disassembly procedures; recalling the operation and check out procedures for a piece of equipment; following equipment turn-on procedures -- emphasis on motor behavior.)

26. Estimates speed: (Concerns the speed of moving objects or materials relative to a fixed point or to other moving objects. For example, the speed of vehicles.)

27. Estimates distances: (Concerns the distance from one location to another. For example, from observer's location to an object on the horizon.)

28. Adopts proper attitude: (Concerns exhibiting a pattern of behavior consistent with an attitude or value; a willingness to perform according to a standard as opposed to skill to perform according to that standard. Integrating or organizing a value or attitude into a pattern of behavior. For example, complying with known safety standards while performing a maintenance procedure on a high voltage power supply.)

OVERT RESPONSES

29. Finger manipulation: (Concerns making finger movements in various types of activities; usually the hand and arm are not involved to any great extent. For example, indexing announced ammunition into computer.)

30. Hand-arm movement: (Concerns the manual control or manipulation of objects through hand or arm movements, which may or may not require continuous visual control; requires coordination of hand-arm movements. For example, pull charging handle of M85 machinegun rearward until bolt locks in place; open breech.)

31. Foot-leg movement: (Concerns the manual control or manipulation of objects through foot or leg movements, which may or may not require continuous visual control; requires coordination of foot-leg movements. For example, lock parking brakes on a tank.)

32. Steers: (Concerns compensatory movements based on feedback from displays; involves estimating changes in positions, velocities, accelerations and a knowledge of display -- control relationships. For example, tank driver following a road.)

33. Tracks: (A perceptual-motor activity involving continuous pursuit of a target or keeping dials at a certain reading; requires smooth muscle coordination patterns -- lack of overcontrol. For example, tank-gunnery target tracking; sonar operator keeping the cursor on a sonar target.)

34. Reports in writing: (Concerns the copying or posting of information for immediate or later use. For example, transcribing a radio message; noting maintenance faults on DA Form 2404.)

35. Reports by talking: (Concerns the oral passage of routine or nonroutine information or facts. For example, announce UP, announce IDENTIFIED.)

36. None: (The task or sub-task has no overt response.)

APPENDIX H

EIGHTEEN TASK SAMPLE USED IN THE PRACTICE RATINGS

EIGHTEEN TASK SAMPLE USED IN THE PRACTICE RATINGS

1. Perform before-operations maintenance checks on hydraulic brake system (Driver).
2. Perform before-operations maintenance checks and services on tank engine and transmission oil levels (Driver).
3. Install the M24 (IR) periscope (Driver).
4. Start tank engine (Driver).
5. Perform during-operations maintenance checks and services on steering, accelerator, shift and brake controls (Driver).
6. Remove the main gun breechblock group (Loader).
7. Disassemble the breechblock (Loader).
8. Perform main gun prepare-to-fire procedures from the Loader's position (Loader).
9. Clear an M219 machinegun (Loader).
10. Load an M219 machinegun (Loader).
11. Prepare tank for boresighting (Loader).
12. Prepare tank for boresighting (Gunner).
13. Boresight Gunner's Telescope (Gunner).
14. Zero an M219 machinegun (Gunner).
15. Boresight rangefinder with the main gun bore axis alined on an aiming point at 1200 meters (Tank Commander).
16. Mount an M85 machinegun in a tank (Tank Commander).
17. Clear an M85 machinegun (Tank Commander).
18. Prepare tank for boresighting (Tank Commander).

APPENDIX I

TWENTY-TWO TASK SAMPLE USED TO VERIFY INTER-RATER RELIABILITY

TWENTY-TWO TASK SAMPLE USED TO VERIFY
INTER-RATER RELIABILITY

1. Perform before-operations maintenance checks on fire extinguishers (Driver).
2. Stop tank engine (Driver).
3. Start tank engine by auxiliary power -- slave start (Driver).
4. Connect track (Driver).
5. Perform after-operations maintenance checks and services on the gun travel lock (Driver).
6. Perform after-operations maintenance checks and services on the tank batteries (Driver).
7. Adjust variable breech operating cam (Loader).
8. Perform emergency closing of main gun breech (Loader).
9. Remove an M219 machinegun from a tank (Loader).
10. Drain replenisher system (Gunner).
11. Operate Gunner's quadrant (Gunner).
12. Apply immediate action in case of main gun failure to fire (Gunner).
13. Acquire ground targets (night) (Tank Commander).
14. Apply immediate action to reduce stoppage of an M85 machinegun (Tank Commander).
15. Gunner fires range card lay to direct fire using Gunner's telescope and coax (stationary/moving).
16. Tank Commander fires nonprecision .50 caliber engagement using the TPI (moving/moving).
17. Tank Commander fires nonprecision coax engagement using the RFI (moving/moving).
18. Tank Commander fires main gun battlesight engagement using the RFD (moving/stationary).
19. Gunner fires main gun battlesight to precision engagement using the GPD (moving/stationary).
20. Gunner fires coax precision engagement using the TEL (moving stationary).
21. Tank Commander fires main gun range card lay to direct fire using the RFD (stationary/stationary).
22. Gunner fires main gun precision engagement using the TEL (stationary/moving).

APPENDIX J

INTER-RATER RELIABILITY STUDIES: COMPUTATION DETAILS AND DISCUSSION OF RESULTS

INTER-RATER RELIABILITY STUDIES: COMPUTATION DETAILS AND DISCUSSION OF RESULTS

COMPUTATION

A phi coefficient was computed for each subset of task descriptors (Stimuli; Tools, Instruments and Controls; Mediating Processes; Overt Responses) as well as the total (across subsets) for each of the 18 tasks both before and after rater discussion. The data for each task were organized into two-by-two bivariate frequency tables for each descriptor subset and for the total. Data were entered in 180 tables (four subsets and total, by 18 tasks, both before and after rater discussion) as follows:

	$R_2 = 0$	$R_2 = 1$
$R_1 = 0$	a	b
$R_1 = 1$	c	d

R_1 = Rater 1

R_2 = Rater 2

where a = number of cells corresponding to task descriptors in a subset that both raters agreed <u>were not</u> included in subtasks of the task.

b = number of cells corresponding to task descriptors in a subset that Rater 1 said "is not" and Rater 2 said "is" included in subtasks of the task.

c = number of cells corresponding to task descriptors in a subset that Rater 1 said "is" and Rater 2 said "is not" included in subtasks of the task.

d = number of cells corresponding to task descriptors in a subset that both raters agreed were included in subtasks of the task.

Figure J.1 is a sample rating sheet for preparing the two-by-two bivariate. frequency table for the Stimuli subset of one of the tasks in the sample. Entries were made as follows:

	$R_2 = 0$	$R_2 = 1$
$R_1 = 0$	26	3
$R_1 = 1$	1	3

$\Sigma = 33$

STIMULI

DRIVER

PERFORM BEFORE-OPERATIONS MAINTENANCE CHECKS ON HYDRAULIC BRAKE SYSTEM	1. Written (textual) material	2. Graphic/tabular material	3. Instrument read-outs	4. Natural environmental features	5. Man-made environmental features	6. Oral command or request	7. Non-verbal sounds	8. Smell (olfaction)	9. Body feel (kinesthesis)	10. Touch	11. Self-initiated
RATER 1											
1. Apply brake and hold for approximately 30 seconds.	1										
2. Observe brake pressure gage and insure that it indicates and maintains 750-900 PSI.	1		1								
3. Note any drop in pressure as a fault on DA Form 2404.											1
RATER 2											
1. Apply brake and hold for approximately 30 seconds.	1									1	
2. Observe brake pressure gage and insure that it indicates and maintains 750-900 PSI.	1		1								
3. Note any drop in pressure as a fault on DA Form 2404.	1		1								

Figure J.1. Sample rating sheet for preparing two-by-two bivariate frequency table.

The sum of the entries in any table is equal to the product of the number of task descriptors in the subset and the number of subtasks in the task. (Eleven task descriptors by three subtasks = 33 entries).

Since relatively few (typically about a third) of the 36 descriptors were judged as characteristic of a given task, we were concerned that inter-rater reliability coefficients would be inflated by the large number of zero-zero agreements. This is a valid concern to the extent that for a given task many discriptors are so totally and obviously irrelevant that a "0" rating requires little intelligent judgment on the part of the raters. To correct for this possibility, phi coefficients were computed using selected descriptors in each case.

The coefficient was computed by first reducing the entries in cell "a" of each bivariate frequency table by the product of the number of task descriptors in any subset irrelevant to a particular task and the number of subtasks in the task. For example, the two-by-two bivariate frequency table for the Stimuli subset of the task in Figure J.1 was as follows:

	$R_2 = 0$	$R_2 = 1$
$R_1 = 0$	5	3
$R_1 = 1$	1	3

Seven task descriptors (graphic/tabular material, natural environmental features, man-made environmental features, oral command or request, non-verbal sounds, smell, and body feel) were considered by both raters irrelevant to the set of subtasks comprising this task; cell "a" was therefore reduced by 21 (7 task descriptors by 3 subtasks). The selected descriptors used to compute the phi coefficient for this subset were written (textual) material, instrument read-outs, touch, and self-initiated. No other cell entries were reduced by this procedure.

All coefficients of inter-rater reliability reported in the following section were computed using the more conservative selected descriptors approach, an approach yielding coefficients that averaged about .055 correlational points less than those based on all descriptors. Results of the two computational approaches are compared in Appendix K.

RESULTS

Effects of Rater Discussion

Inter-rater reliabilities for the 18 practice tasks are shown by descriptor subset and rating period (before vs. after discussion) in Table J.1. The coefficients in the body of the table show considerable variation, and since many are based on fewer than 20 observations, interpretations at the task-by-descriptor level probably are not useful. At the total task level, however, the correlations are more stable. All but two of the 36 rater agreement coefficients by task (right-hand column of Table J.1) were significant at the .05 level. The before-discussion reliabilities for Tasks 5 and 18, which were .20 and .12 respectively, were not significant.[1]

The effects of rater practice and discussion can be seen in the bottom row of Table J.1. Total (across-descriptor) inter-rater reliability increased after discussion, as did the reliabilities for each descriptor category. The increase from .58 to .72 in total inter-rater reliability was significant at the .05 level.[2] The increase in the reliabilities for all but the Stimuli category of descriptors also were significant at the .05 level.[2]

Differences in reliability as a function of descriptor category also are worth noting. Inter-rater reliability was highest for the Overt Response category both before and after discussion, and was lowest

---

[1] $[\phi = .20] < [r_{.95} \text{ with } 28 \text{ df} = .31]$

$[\phi = .12] < [r_{.95} \text{ with } 46 \text{ df} = .24]$

[2] The difference was evaluated statistically using a chi-square type analysis of the transformed Fisher's z correlation (Hays, 1967, p. 532).

Table J.1

INTER-RATER RELIABILITIES (Ø) FOR THE 18-TASK SAMPLE
BEFORE AND AFTER RATER DISCUSSION

		TASK DESCRIPTOR SUBSETS				
TASK	RATING PERIOD	STIMULI (N)	TOOLS, INSTMTS CONTROLS (N)	MEDIATING PROCESSES (N)	OVERT RESPONSES (N)	TOTAL (N)
1	BEFORE	.845 (12)	1.00 (3)	.293 (12)	1.00 (6)	.694 (33)
	AFTER	.550 (9)	.671 (11)	1.00 (3)	1.00 (9)	.778 (32)
2	BEFORE	.633 (21)	.671 (21)	-.158 (21)	.867 (14)	.518 (77)
	AFTER	.848 (14)	.919 (28)	-.221 (28)	1.00 (14)	.606 (84)
3	BEFORE	1.00 (9)	.000 (9)	NR[1] (0)	.892 (18)	.835 (36)
	AFTER	.000 (9)	.478 (9)	NR (0)	.894 (18)	.717 (36)
4	BEFORE	.501 (56)	.576 (42)	.129 (70)	.791 (42)	.562 (210)
	AFTER	.504 (56)	.696 (42)	.128 (56)	.930 (28)	.643 (182)
5	BEFORE	.000 (4)	.577 (4)	-.255 (12)	.500 (10)	.200 (30)
	AFTER	1.00 (4)	.577 (4)	.447 (6)	.816 (10)	.707 (24)
6	BEFORE	.752 (38)	.623 (57)	.716 (57)	.854 (38)	.745 (190)
	AFTER	.881 (38)	.936 (76)	.255 (76)	.948 (38)	.841 (228)
7	BEFORE	NR (0)	1.00 (6)	NR (0)	.674 (12)	.886 (18)
	AFTER	NR (0)	1.00 (6)	.000 (12)	.357 (12)	.591 (30)
8	BEFORE	.747 (72)	.511 (72)	.190 (72)	.527 (54)	.552 (270)
	AFTER	.715 (90)	.851 (90)	.753 (72)	.841 (54)	.805 (306)
9	BEFORE	.804 (36)	1.00 (12)	.469 (34)	.500 (36)	.688 (118)
	AFTER	.217 (24)	.582 (36)	.692 (24)	.942 (36)	.706 (120)
10	BEFORE	.645 (50)	1.00 (10)	-.050 (30)	1.00 (20)	.831 (110)
	AFTER	.608 (20)	.614 (30)	.464 (30)	.302 (20)	.563 (100)
11	BEFORE	.000 (12)	.756 (9)	.632 (0)	.632 (6)	.644 (27)
	AFTER	1.00 (6)	1.00 (6)	1.00 (3)	.000 (6)	1.00 (21)
12	BEFORE	.258 (28)	-.250 (21)	NR (14)	.333 (28)	.189 (91)
	AFTER	.632 (28)	1.00 (28)	.000 (21)	1.00 (28)	.806 (105)

Table J.1 (Continued)

13	BEFORE	-.121 (55)	.471 (44)	.000 (66)	.278 (55)	.159 (220)
	AFTER	.806 (55)	.533 (33)	.583 (55)	.913 (44)	.723 (187)
14	BEFORE	.129 (39)	.619 (43)	.174 (26)	.741 (39)	.386 (147)
	AFTER	.471 (26)	.571 (39)	.186 (39)	.939 (52)	.566 (156)
15	BEFORE	1.00 (0)	.621 (0)	.000 (8)	.617 (24)	.648 (32)
	AFTER	.659 (0)	.707 (16)	1.00 (8)	.872 (8)	.818 (32)
16	BEFORE	NR (18)	NR (18)	1.00 (18)	.730 (18)	.778 (72)
	AFTER	NR (27)	.745 (27)	1.00 (18)	.000 (18)	.881 (90)
17	BEFORE	.791 (3)	.614 (9)	.686 (6)	.342 (6)	.614 (24)
	AFTER	.250 (3)	.500 (9)	.000 (3)	.892 (6)	.626 (21)
18	BEFORE	.000 (12)	.745 (8)	-.135 (12)	-.041 (16)	.124 (48)
	AFTER	.816 (12)	.837 (12)	1.00 (8)	.618 (16)	.778 (48)
ALL TASK	BEFORE	.578 (465)	.610 (388)	.221 (458)	.661 (442)	.576 (1753)
	AFTER	.634 (421)	.744 (502)	.438 (462)	.859 (417)	.728 (1802)

[1]NR = NONE RATED

for Mediating Processes. The rank-order of reliabilities for the descriptor categories was the same before and after discussion.

Verification Study

As noted earlier, 22 of the 208 M60A1 tasks that were not rated in the practice session were rated using the same methods and raters as were used for the 18 practice tasks. The ratings of the 22-task sample were compared to the second-round ratings of the 18-task sample, as a means of verifying the level of inter-rater reliability attained in the final round of ratings for the 18 practice tasks, and as a check on the independence of the final ratings of the 18 practice tasks.

Phi coefficients, computed as in the practice ratings, are presented in Table J.2. Here it can be seen that the rank-order of the reliabilities for the four descriptor categories is the same as the before-and-after rank-orders in the practice ratings. Overt Responses and Mediating Processes were highest and lowest, respectively.

Inter-rater reliabilities for the two samples are presented in Table J.3, where it can be seen that the reliabilities were consistently lower for the 22-task sample than for the 18-task sample. The differences between the reliabilities for the two samples are significant (.05 level) for each descriptor category except Mediating Processes, and for the total across descriptors.

Combined reliabilities also are shown in Table J.3 (bottom row). The combined coefficients are not the means for the two samples. Rather the coefficients were obtained by treating the two samples as one 40-task sample, and computing five separate phis: one for each of the four descriptor categories, and one for the total across descriptors. The overall reliability for the combined sample approached .70, with Overt Responses and Mediating Processes once again ranking highest and lowest.

Table J.2

INTER-RATER RELIABILITIES (Ø) FOR THE 22-TASK SAMPLE

	TASK DESCRIPTOR SUBSETS				
TASK	STIMULI (N)	TOOLS, INSTMTS CONTROLS (N)	MEDIATING PROCESSES (N)	OVERT RESPONSES (N)	TOTAL (N)
1	.478 (9)	1.00 (3)	.250 (6)	.800 (9)	.586 (27)
2	.556 (12)	.214 (18)	NR*	1.00 (18)	.596 (48)
3	.805 (39)	.709 (65)	.185 (39)	.856 (26)	.675 (169)
4	NR	.300 (40)	-.062 (30)	.790 (30)	.520 (100)
5	.250 (6)	1.00 (2)	.707 (6)	.707 (6)	.583 (20)
6	.057 (33)	.588 (22)	.160 (33)	.866 (33)	.500 (121)
7	NR	1.00 (6)	NR	.333 (6)	.667 (12)
8	NR	.577 (8)	.000 (4)	1.00 (8)	.704 (20)
9	NR	.576 (14)	NR	.745 (14)	.710 (28)
10	1.00 (8)	.408 (12)	.000 (4)	.000 (4)	.624 (28)
11	-.408 (15)	.133 (45)	-.163 (60)	.519 (60)	.191 (180)
12	1.00 (24)	.367 (36)	.000 (12)	.507 (36)	.590 (108)
13	.200 (15)	.000 (5)	-.038 (35)	.166 (10)	.129 (65)
14	.490 (48)	.546 (64)	.194 (48)	.626 (32)	.553 (192)
15	.800 (145)	.937 (87)	.684 (116)	.865 (145)	.845 (493)
16	.324 (33)	.722 (33)	.432 (44)	.714 (66)	.589 (176)
17	.452 (72)	.756 (54)	.390 (90)	.704 (108)	.604 (324)
18	.455 (80)	.770 (48)	.827 (80)	.916 (80)	.762 (288)
19	.543 (125)	.859 (75)	.718 (125)	.867 (125)	.758 (450)
20	.620 (110)	.744 (66)	.642 (110)	.846 (88)	.737 (374)
21	.538 (150)	.903 (75)	.571 (125)	.916 (125)	.751 (475)
22	.580 (138)	.662 (69)	.708 (161)	.752 (138)	.682 (506)
ALL TASKS	.550 (1062)	.671 (847)	.493 (1128)	.779 (1167)	.662 (4204)

* NR = NONE RATED

Table J.3

INTER-RATER RELIABILITIES (Ø) FOR THE 18-TASK (SECOND-ROUND) AND 22-TASK SAMPLES

	STIMULI	TOOLS, INSTS., AND CONTROLS	MEDIATING PROCESSES	OVERT RESPONSES	ALL DESCRIPTORS
18-TASK SAMPLE	.634	.744	.438	.859	.729
22-TASK SAMPLE	.550	.671	.493	.779	.662
BOTH SAMPLES	.573	.697	.478	.804	.682

DISCUSSION

The data from the practice ratings present little interpretive difficulty. Increases in reliability after practice and discussion were observed across descriptors, and in each of the four descriptor categories. The increases were significant for inter-rater reliability across descriptors and for three of the four descriptor categories. The benefit of practice and discussion on inter-rater reliability seems unequivocal.

Interpreting the results of the Verification Study is less straightforward. Inter-rater reliabilities for the 22-task sample were significantly lower overall and in three of the four descriptor categories than were inter-rater reliabilities for the second-round ratings of the 18-task sample. One might be inclined therefore to conclude that the practice effect, while dramatic, is highly specific to the sample of tasks being rated. The tenability of this conclusion may be examined by comparing inter-rater reliabilities for the 22-task sample and for the first-round ratings of the 18-task sample. If the practice effect were specific to the sample of tasks being rated, then no differences would be expected between inter-rater reliabilities for the ratings of the 22-task sample and the first-round ratings of the 18-task sample. The two sets of ratings are presented in Table J.4. Increases in reliability can be seen across descriptors, and in three of the four descriptor categories. All increases were significant. (The decrease in the Stimuli category was not significant.) It appears then that the practice effect has both specific and general components: inter-rater reliability increased significantly when the 18-task sample was re-rated and when the 22-task sample was rated for the first time. That inter-rater reliability was significantly lower for the 22-task sample than for the second-round ratings of the 18-task sample simply suggests that the practice effect is stronger when identical tasks are rated and then re-rated, than when the practice sample is different from the sample that is rated for record. The important point is not that practice affected inter-rater reliability differently for the two samples, but that significant increases in

Table J.4

INTER-RATER RELIABILITIES (Ø) FOR THE 18-TASK (FIRST-ROUND) AND 22-TASK SAMPLES

	STIMULI	TOOLS, INSTS., AND CONTROLS	MEDIATING PROCESSES	OVERT RESPONSES	ALL DESCRIPTORS
18-TASK SAMPLE	.578	.610	.221	.661	.576
22-TASK SAMPLE	.550	.671	.493	.779	.662

inter-rater reliability occurred in both cases. The overall reliability was about .70 in both cases, and was .68 for the combined sample. The coefficients are far in excess of chance expectancy, and are estimates of the inter-rater reliability for all tasks rated after the practice session.

Inherent differences in the difficulty with which tasks may be characterized by each descriptor subset were suggested by the stability of the rank-orders of reliabilities for the descriptor categories in the practice ratings and in the Verification Study. Inter-rater reliability was invariably highest for Overt Responses, probably because descriptors in this category required little definition beyond naming, and were therefore easity judged as required or not required in task performance. The subset for Tools, Instruments and Controls yielded somewhat lower indexes of agreement; the raters disagreed mainly on the use of fixed and variable controls, and on common and special hand tools. Ready access to tanks, as a means of verifying information obtained from technical manuals and experts, would have eliminated many of these disagreements.

Inter-rater reliability for Stimuli was depressed because of fairly consistent disagreement between raters in choosing either self-initiated or oral command/request descriptors. Many of these disagreements probably could have been eliminated by pinpointing their sources early in the rating process, and increasing the precision of the descriptor definitions.

Mediating Processes consistently yielded the lowest inter-rater reliability. The descriptors in this category were not mutually exclusive, not easily defined or remembered, and offered no external criteria against which the raters could evaluate the validity of their judgments. More precise descriptor definitions and additional rater practice might have improved reliability here.

CONCLUSIONS

Among the conclusions that can be drawn from the inter-rater reliability studies are:

1. Inter-rater reliability increased significantly with practice and discussion, irrespective of whether the tasks rated for record were the same as or different from the tanks rated for practice.

2. Overall inter-rater reliabilities for the tasks rated after practice were about .70.

3. Inter-rater reliability varied consistently as a function of descriptor subsets. Reliability was invariably highest for Overt Responses and lowest for Mediating Processes.

4. Increases in inter-rater reliability greater than those obtained in the present studies probably could have been achieved with:

    A. Increased precision and clarity of the descriptor definitions.

    B. More practice.

    C. More access to operational equipment, as a means of verifying information obtained from technical manuals and experts.

APPENDIX K

PHI COEFFICIENTS BASED ON ALL DESCRIPTORS COMPARED TO PHI COEFFICIENTS BASED ON SELECTED DESCRIPTORS

PHI COEFFICIENTS BASED ON ALL DESCRIPTORS COMPARED TO PHI COEFFICIENTS BASED ON SELECTED DESCRIPTORS

EIGHTEEN TASK SAMPLE
(COMBINED PHI FOR BEFORE AND AFTER RATINGS)

	DESCRIPTOR SUBSETS				TOTAL
	STIMULI	TOOLS, INST. CONTROLS	MEDIATING PROCESSES	OVERT RESPONSES	
ALL DESCRIPTORS	.665	.772	.397	.845	.717
SELECTED DESCRIPTORS	.605	.691	.334	.776	.659

TWENTY-TWO TASK SAMPLE

	DESCRIPTOR SUBSETS				TOTAL
	STIMULI	TOOLS, INST. CONTROLS	MEDIATING PROCESSES	OVERT RESPONSES	
ALL DESCRIPTORS	.617	.720	.535	.815	.713
SELECTED DESCRIPTORS	.550	.671	.493	.779	.662

APPENDIX L

CLUSTER ANALYSIS PROCEDURES

CLUSTER ANALYSIS PROCEDURES

Each cluster analysis began by calculating the "behavioral distance" between every pair of tasks. Many distance measures have been reported in the literature, but for the one-zero data in the task by task-descriptor matrix, most of the measures are equivalent. The Simple Matching Coefficient (SMC) was used to measure behavioral distance in the present analyses. The SMC measures distance by the proportion of task descriptors that is identical between each pair of tasks. Thus for two tasks that have exactly the same values on 12 of the 36 descriptors, the intertask distance is 12/36 or .33.

Two clustering algorithms which employ the SMC were considered. One of these, the Average Distance Amalgamation algorithm,[1] has long been used to form clusters with the kind of data available, but requires an assumption that the 36 task descriptors are orthogonal. Since this assumption seemed questionable, another algorithm which does not require the orthogonality assumption, the Direct Clustering algorithm,[2,3] was used.

Use of the SMC produces a matrix that shows the behavioral distance between every pair of tasks. Tasks that are "close together" in behavioral distance form the task clusters or skills. The process is amalgative, in that the two closest tasks form the seed for the first cluster. Nearby tasks are incorporated into this cluster until a task is found that is too far away; this task then forms the seed of a new cluster. Clusters amalgamate similarly. In the first pass of the analysis, each task forms a cluster. Successive passes produce fewer and fewer clusters, each containing more and more tasks, until on the final pass all tasks are included in a single cluster. Selecting passes and clusters within passes is driven by the purposes for doing so.

---

[1]Dixon, W.J., op. cit., 1975.

[2]Hartigan, J.A., op. cit., 1972.

[3]Dixon, W.J., op. cit., 1975.

SELECTING PASSES AND CLUSTERS

The task-joining sequences for each of the four duty positions are presented in Figures L.1, L.2, L.3, and L.4. The clusters that formed in each pass are indicated by brackets; the clusters that were selected to represent skills are indicated by heavy lines. The tasks comprising each skill are presented by duty position in Appendix B.

The procedure for selecting passes and clusters is constrained by the requirement that the integrity of clusters be maintained. One examines the clusters as they form larger clusters from pass to pass. Since (by definition) any cluster contains tasks grouped according to similar task descriptors, a criterion other than similar descriptors is needed for selecting clusters. The criterion that was used was to try to find the smallest number of clusters that were:

1. Dissimilar operationally from one another.
2. Each comprised of functionally or operationally related tasks.

After examining the clusters, it became apparent that the criterion could not be rigorously applied in all cases. Some compromises were required.

When the tasks comprising a cluster described similar mission operations, we selected that cluster and gave it a title in terms of its mission characteristics. When the tasks did not describe similar mission operations, we used the clusters from the preceding pass unless they numbered more than four. When there were more than four clusters in the preceding pass, the non-similar task cluster was used and described in mission-operation terms which defined most of the tasks in the cluster. These clusters are indicated in Appendix B by an asterisk. Sometimes two or three dissimilar tasks formed a cluster during Pass 1 and remained a unique cluster until the final pass. When this happened, the integrity of the cluster was maintained. An example is Cluster 9 for the Gunner,

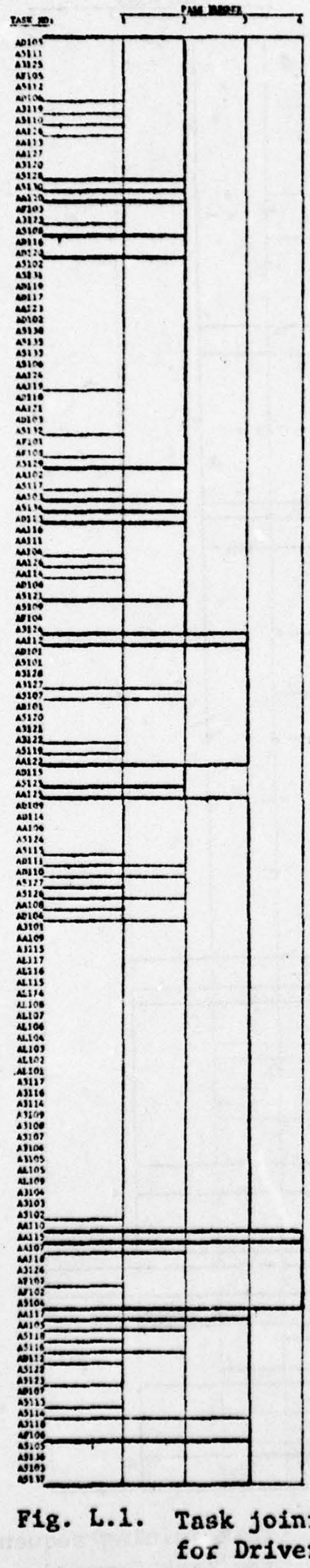


Fig. L.1. Task joining sequence for Driver tasks.

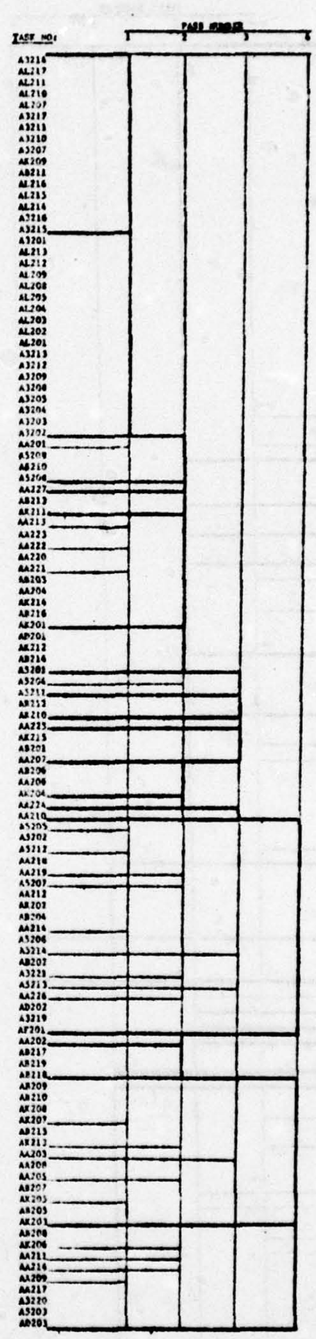


Fig. L.2. Task joining sequence for Loader tasks.

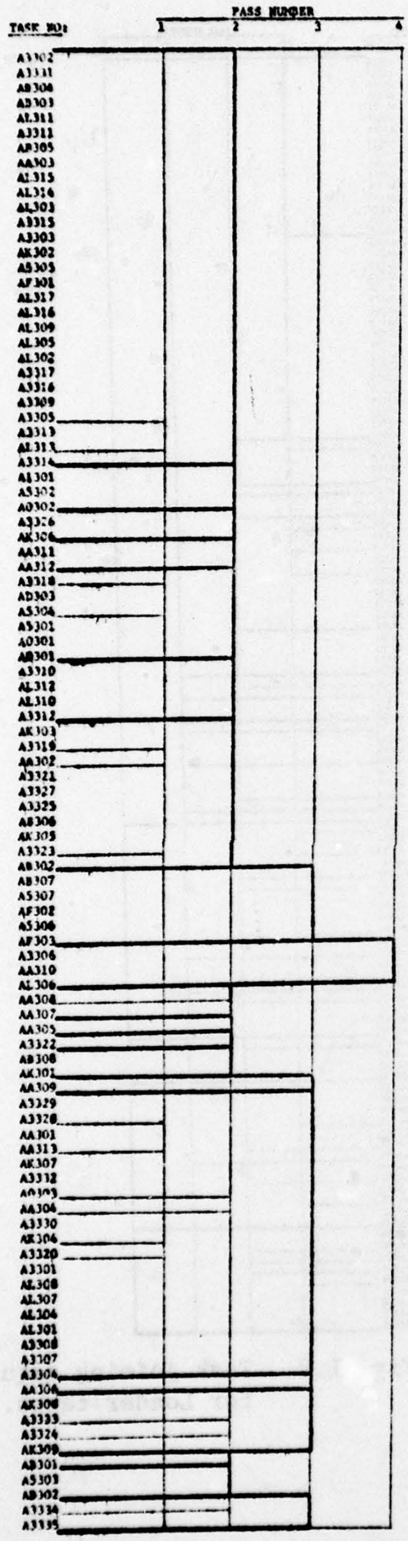


Fig. L.3. Task joining sequence for Gunner tasks.

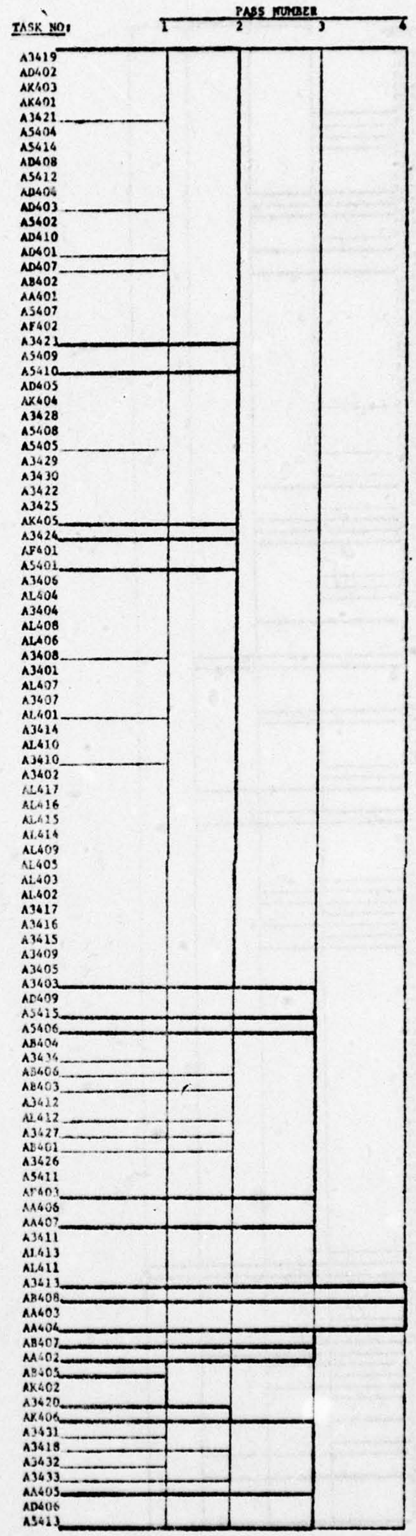


Fig. L.4. Task joining sequence for Tank Commander tasks.

"Assist in Night .50 Caliber Engagements," which is a three-task cluster. Two of the tasks (A3306 and AL306) pertain to assisting in a .50 caliber engagement, and the third task (AA310) is an azimuth indicator task. They formed a cluster during Pass 1 and remained together in all successive passes.

In two cases -- Cluster 5 for the Gunner and Cluster 9 for the Tank Commander -- the clusters were divided into two clusters to make them more homogeneous in terms of mission operations.

DESCRIBING THE SKILLS

Skill descriptions were written after the clusters were selected and named. For example, the skill description for Tank Commander's Cluster 1, "Operate Weapon Systems," was:

> Performs fixed procedure, finger-hand-arm manipulation of various controls in voluntary response to man-made environmental features, non-verbal sounds, or touch, by recalling facts, detecting or classifying information.

The method for describing the skills was generally to mention overt responses first; then the tools, instruments, and controls; next, the stimuli associated with the responses; and finally, the mediating process. The formula was: "Performs [OVERT RESPONSE(S)] of [TOOLS, INSTRUMENTS, AND CONTROLS], in response to [STIMULI] by [MEDIATING PROCESSES]." Application of the formula was by no means hard and fast. Variations in the descriptions resulted from using the following guidelines:

1. Task descriptors that appeared in greater than 50 percent of the tasks in a cluster were mentioned.
2. Task descriptors that appeared in 30 to 50 percent of the tasks in a cluster were mentioned, preceded by "sometimes."
3. The task descriptor "recalls set procedures" was placed after "Performs" and changed to "fixed procedure."
4. When all the controls occurred, the words "various controls" were used.

5. The task descriptor "steers" was changed to "continuous manipulation"; "tracks" was changed to "compensatory manipulation," and placed after "Performs."

6. When "foot-leg movement" occurred with "finger manipulation," "hand-arm movement," or both, "multi-limb manipulation" was used.

7. When both "oral command or request" and "reports by talking" occurred, "communicates orally" was used and placed before "Performs."

8. When "reports by talking," "reports in writing" or both occurred, each was placed after the mediating processes.

9. The task descriptor "self-initiated" was changed to "voluntary response."

APPENDIX M

LEARNING AND EVALUATION DIFFICULTY STUDY

LEARNING AND EVALUATION DIFFICULTY STUDY

This part of Task 1 was aimed at obtaining estimates of the relative difficulty of learning and evaluating the skills identified in the cluster analysis. The estimates were derived from the judgments of members of the project staff, who rated the task descriptors in terms of the relative training difficulty and the relative evaluation difficulty for the domain of tank crew behavior associated with each descriptor. Difficulty estimates for each skill were made by assigning the descriptor ratings to the modal descriptor pattern for each skill.

Descriptors rather than skills were rated for several reasons. The main reason was that rating the descriptors provides a set of stable scores, which in turn provide flexibility that might be needed later in the project. If, for example, learning or evaluation-difficulty scores at the task level are desired, they are easily obtained: one simply examines the descriptor pattern for the task on the one hand, and the descriptor scores on the other. A task rating is derived by combining the scores appropriate to the descriptor pattern of the task. Similarly, if task clusters are combined or further divided later, it will not be necessary to conduct new studies to obtain learning- and evaluation-difficulty scores for the new clusters. The descriptor patterns for the new clusters can be examined and new ratings derived by combining the descriptor scores that correspond to the descriptor patterns.

Another reason for not rating the skills directly was that the skills are global, and thus invite unreliability in ratings. If exemplar tasks are given the rater for each skill, then the risk is that the ratings will be made of the exemplar tasks only, and not of the skill as a whole. If raters are given the population of tasks for each skill, unreliability is once again invited: some raters will focus on one part of the population, and others on other parts. If raters are given only the skill title and description with no reference to tasks, the problem remains. Raters will invent their own exemplar tasks, which may differ from rater to rater. The consequence is degraded inter-rater reliability, because raters are rating "different things."

Use of a partial paired comparison study, similar or identical in all essentials to the criticality study described earlier, also was considered and abandoned. One reason was that at least two such studies would be required -- one for learning difficulty and another for evaluation difficulty. Tabulating and analyzing paired-comparison studies would have placed demands on project resources that could not have been met.

RATERS

Five members of the project staff, two of whom had performed the original ratings of the tasks in terms of the 36 descriptors, and all of whom were familiar with the project purposes and proposed methodology, performed the difficulty ratings.

PROCEDURE

A list of the 36 descriptors with four descriptors deleted was given to each rater, along with the descriptor definitions that appear in Appendix G. The four deleted descriptors were ones that were used by neither of the two raters in the original task characterization: "smell" in the Stimuli subset; "none" in the Tools, Instruments, and Controls subset; "identifies symbols" in the Mediating Process subset; and "none" in the Overt Responses subset.

The raters were asked to assign three numbers from an absolute scale of one (extremely easy to learn or evaluate) to 50 (extremely difficult to learn or evaluate) to the domain of tank crew behavior associated with each descriptor. The three ratings of each descriptor were to represent:

1. Learning difficulty.
2. "Hands-on" performance evaluation difficulty (where test validity is not a problem).
3. Difficulty of evaluation by any means, while maintaining acceptable validity, and trading off validity against economy.

Additional details of the instructions to the raters may be found in Appendix N.

After the raters had considered the descriptors in terms of the three factors, they discussed their interpretations of the descriptors, and were permitted to adjust their ratings of difficulty. Only the second set of evaluation difficulty ratings, representing difficulty of any means of testing, including full-performance testing, were used to determine skill evaluation difficulty; the full-performance evaluation difficulty ratings were requested so that the raters would first assign ceiling values to each descriptor's difficulty. The ratings of difficulty of evaluating by any means would then be the same as or lower than those of full-performance testing, depending on the feasibility of other means and the sacrifice in validity.

RESULTS

Difficulty Scales

The values assigned to the 32 descriptors on learning and evaluation difficulty were averaged across raters, and the mean values were used in computing the skill difficulties. For the modal pattern of descriptors for each skill, the difficulty values of those descriptors were summed separately for learning and evaluation difficulty. The skill learning difficulties (sums ranged from 87 to 456, and the evaluation difficulties ranged from 58 to 287. Although these values represent not only the separate difficulty values assigned to individual descriptors, but also the number of descriptors comprising each skill, it was felt that the skill difficulty as an additive function of difficulty of the descriptors would be reflected better by the sum than by the mean. The sums were converted to standardized scales for learning and evaluation difficulty, each with a mean of 5.00 and standard deviation of 1.00, the same standard scale as was used for criticality ratings. The standardized scale values for each skill were presented in Tables 4 through 7.

Reliability

Inter-rater reliability was estimated by an analysis of variance of the rater by descriptor data matrix.[1] Intraclass correlations were .76 for learning difficulty and .88 for evaluation difficulty, indicating fairly high reliability of the average of the five sets of ratings. (Each coefficient indicates the hypothetical correlation that would obtain between the average ratings for this set of five raters and those from another random sample of five raters.) If it is assumed, however, that the raters differed systematically in their frames of reference for judging the descriptors, then the reported correlations are underestimates of inter-rater reliability. When the data are corrected for differences among rater means, reliability of the mean ratings are .85 for learning difficulty, and .89 for evaluation difficulty.

[1] Winer, B.J., op. cit., 1962.

APPENDIX N

INSTRUCTIONS TO RATERS FOR
THE LEARNING AND EVALUATION
DIFFICULTY STUDIES

INSTRUCTIONS TO RATERS FOR THE
LEARNING AND EVALUATION DIFFICULTY STUDIES

A list of 32 behavioral descriptors is attached, along with a set of definitions of the descriptors.

We need to get your judgments about the difficulty of learning, and the difficulty of evaluating, behavior associated with each descriptor.

The difficulty judgments are to be made with respect to the entire domain of tank crew behavior. Thus, if you're making a judgment about the learning difficulty associated with the descriptor "Graphic/tabular material," you should think in terms of the domain of tank crew behaviors that involve using or responding to graphic or tabular materials. Then the question to ask yourself is "How difficult would it be to learn the behavior in this domain, relative to learning the behaviors in the domains associated with the other discriptors?"

Learning difficulty is defined as the amount of time, practice, or trials to criterion that would be required to attain proficiency in the domain of behavior associated with each descriptor.

Evaluation difficulty is less straight-forward. Here we'd like two separate sets of ratings. The first set is concerned exclusively with "hands-on" performance evaluation, where test validity is assumed not to be a problem. That is, if we had our choice among high-fidelity performance tests, then we could assume that validity is acceptable. The judgments about evaluation difficulty therefore would be made on the basis of considerations other than validity. The judgments probably reduce to considerations of economy: Given that the "hands-on" performance tests will yield acceptable validity, which of the tank crew behaviors are more or less expensive to test in the "hands-on," full-performance mode? Factors that come into play here are, as you know, equipment costs and scarcity, requirements for scarce terrain, amounts of time required for testing, difficulty of standardization, and numbers and kinds of personnel required to develop and administer the tests. Ultimately then your judgments here will reduce to "How difficult (expensive) would it be to evaluate the behavior in a 'hands-on' mode?" Or, "How expensive would it be to conduct a 'hands-on' performance test?"

In the second set of evaluation difficulty ratings we are not concerned exclusively with the "hands-on" performance setting. Rather, we would like your judgments as to how difficult it would be to evaluate the behavior by any means, and still maintain what in your view would be acceptable test validity. If in your view an inexpensive paper-and-pencil test could be used to measure with acceptable validity the behavior associated with one of the 32 descriptors, then the descriptor would get a lower evaluation difficulty rating than would a descriptor that would require a more expensive full-performance or simulator-based test. Here you are being asked to trade off economy and validity in evaluating the behavior associated with each descriptor.

To summarize: you're being asked for three sets of ratings:

(1) Learning difficulty.

(2) "Hands-on" performance evaluation difficulty (where validity is not a problem).

(3) Difficulty of evaluation by any means, while maintaining acceptable validity, and trading off validity against economy.

Please assign three numbers to each descriptor -- one for learning difficulty, the other two for the two kinds of evaluation difficuly discussed above. The numbers must be between one and 50, where 1 = extremely easy to learn, or extremely easy to evaluate, and 50 = extremely difficult to learn or evaluate. Don't try to do all three sets of judgments at the same time. Do them individually.

Use the definitions liberally. Don't assume that the descriptors are self-explanatory. Many are not. Work independently of the other raters. Take as much time as you need.