Technical Report 584

JOB SAMPLE TESTS AS PREDICTORS OF M1 GUNNERY PERFORMANCE

David W. Biers and Daniel W. Sauer
SYSTEMS RESEARCH LABORATORIES, INC.

ARI FIELD UNIT AT FORT KNOX, KENTUCKY

U. S. Army
Research Institute for the Behavioral and Social Sciences

December 1982

Approved for public release; distribution unlimited.

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

L. NEALE COSBY
Colonel, IN
Commander

EDGAR M. JOHNSON
Technical Director

Research accomplished under contract
for the Department of the Army

Systems Research Laboratories, Inc.

NOTICES

<u>DISTRIBUTION:</u> 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-POT, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.

<u>FINAL DISPOSITION:</u> 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.

<u>NOTE:</u> 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.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report 584	2. GOVT ACCESSION NO.: A139433
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): JOB SAMPLE TESTS AS PREDICTORS OF M1 GUNNERY PERFORMANCE	5. TYPE OF REPORT & PERIOD COVERED: Final Report, November 1980 to September 1982
	6. PERFORMING ORG. REPORT NUMBER: SRL-6639-1
7. AUTHOR(s): David W. Biers and Daniel W. Sauer	8. CONTRACT OR GRANT NUMBER(s): MDA903-81-C-0031
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Systems Research Laboratories, Inc. 2800 Indian Ripple Road Dayton, OH 45440	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q263731A792
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue, Alexandria, VA 22333	12. REPORT DATE: December 1982
	13. NUMBER OF PAGES: 116
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

Contracting Officer's Technical Representative was Ms. Barbara A. Black. Appendixes A, B, C, D, and E were published separately as Research Note 83-32.

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

Personnel selection	M1 tank	M1 computer control panel
Aptitude measurement	Tank commander	Gun laying
Job sample test	Gunner	Target engagement
Tank crewmen	Loader	Tracking
M60A1 tank	Driver	

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

The objectives of Phase I of this research were to (1) develop an aptitude measurement methodology which could be used to design job sample tests for armor crewmen; (2) apply the methodology to develop job sample tests; and (3) administer the job sample tests to armor crewmen and analyze the test data. Phase II, reported separately, included analyses of the predicted validity of the job sample tests.

(Continued)

Item 20 (Continued)

A five-stage methodology for job sample test design was developed. Stages included task identification; task prioritization; job sample dimensional analyses; trade-off analyses; and detailed job sample test development. Seven job sample tests, three computer-based and four hands-on tests, were developed using the methodology. They were Operate Computer Panel, Computer Tracking, Computer Target Engagement, Tank Commander Decision Making, Hands-On Gun Laying, Hands-On Tracking, and Hands-On Target Engagement. Tests were administered to armor crewmen stationed in Europe. The analysis of test data indicated a low degree of intercorrelation among job sample tests which suggested that they were measuring different gunnery behaviors.

Crew experience, in general, was not related to job sample test performance. There was generally good evidence for construct validity of the tests. Although no post predictor criteria were available for Phase I of the research, the results of regression analyses indicate that linear combinations of the job sample test measures account for a very high proportion of the variability in a crew's past success at Annual Qualifications.

Appendixes A, B, C, D, and E were published separately as Research Note 83-32.

Technical Report 584

JOB SAMPLE TESTS AS PREDICTORS OF M1 GUNNERY PERFORMANCE

David W. Biers and Daniel W. Sauer
SYSTEMS RESEARCH LABORATORIES, INC.

Submitted by:
Donald F. Haggard, Chief
ARI FIELD UNIT AT FORT KNOX, KENTUCKY

Approved by:
Harold F. O'Neil, Jr., Director
TRAINING RESEARCH LABORATORY

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

December 1982

Army Project Number 2Q263731A792

Manpower and Personnel

Approved for public release; distribution unlimited.

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The Fort Knox Field Unit has conducted research in the area of Armor personnel selection and assignment for the past several years. Validation research efforts have involved both paper-and-pencil and job sample performance predictors in an attempt to provide the Army with improved methods for selecting and assigning crewmen.

Decisions to fill vacated tank commander and gunner positions are normally made at the operational unit level based on the commander's professional judgment and the soldier's time in grade. Additional information upon which to base these decisions could be provided by job sample tests administered at the unit level using available resources.

This report describes a methodology for the development of job sample tests and demonstrates the application of that methodology to the jobs of M1 tank commander and gunner. The results of this research provide a detailed procedure whereby job sample tests can be constructed and evaluated for numerous Army MOSs.

EDGAR M. JOHNSON
Technical Director

JOB SAMPLE TESTS AS PREDICTORS OF M1 GUNNERY PERFORMANCE

EXECUTIVE SUMMARY

Requirement:

Decisions to select and assign armor crewmen to the crew positions of gunner and tank commander are usually made at the operational unit level. Job sample tests administered at the unit level using unit resources can provide additional information upon which to base a selection and assignment decision. The objectives of this research were to (1) develop an aptitude measurement methodology which could be used to design job sample tests for armor crewmen; (2) apply the methodology to develop job sample tests; and (3) administer the job sample tests to armor crewmen and analyze the test data.

Procedure:

An aptitude measurement methodology was developed to design job sample tests. The methodology included five stages: task identification; task prioritization; job sample dimensional analyses to identify potential means of representing task stimuli, task devices, and eliciting task behaviors; trade-off analyses to select the job sample tests which can be constructed to meet operational schedules, equipment availability, and funding constraints; and detailed job sample test development. The methodology was used to develop seven job sample tests. The three microcomputer-based tests were: Operate Computer Panel, Computer Tracking, and Computer Target Engagement. The four hands-on tests were: Tank Commander Decision Making, Hands-On Gun Laying, Hands-On Tracking, and Hands-On Target Engagement.

The seven job sample tests were administered to M60A1 crewmen stationed in Europe. Crewmembers' biographical data, two ASVAB test scores, and gunnery qualification data were also obtained. Phase I data analyses focused on the biographical data, the relationships within and among job sample tests, and relationships between job sample tests and tank gunnery qualification data. Phase II data analyses focused on the relationships between job sample test performance and M1 transition training performance and M1 gunnery qualifications. Phase II analyses are contained in a separate report.

Findings:

A low degree of intercorrelation was found among the job sample test measures indicating little overlap or redundancy among the measures. This finding suggests that the tests were measuring different behaviors. When combined in a multiple regression equation, these measures should have a greater potential for predicting the criteria variables in Phase II of this research.

Analyses indicated that experience, in general, was not related to job sample test performance. Only in the Hands-On Gun Laying test did the more experienced tank commanders perform better than the less experienced gunners. The lack of statistical evidence makes tenable the assumption that the job sample tests were measuring aptitudes.

Despite the low and nonsignificant correlations, the predicted pattern of relationships between job sample tests emerged in the majority of cases. For two job samples in which relationships were definitely expected (i.e., Computer Tracking and Hands-On Tracking), the predicted pattern emerged and significance was obtained, but for gunners only.

In the absence of any post predictor criteria data, there was also good evidence for the criterion-related validity of the job samples. Results of regression analyses indicated that linear combinations of predictor measures across job samples account for a very high proportion of the variability in past success at Annual Qualifications.

Utilization of Findings:

The five-stage aptitude measurement methodology can guide design of job sample tests, particularly where tasks require a high degree of man-machine interface. The validity of the seven job sample test measures will be established in Phase II. It does appear that, combined in a test battery, the tests have potential for predicting gunnery performance for tank commanders and gunners.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

Section 1

INTRODUCTION

The M1 main battle tank is a totally new tank incorporating advanced technologies in virtually every system. These technologies have contributed to significant improvements in the mobility, survivability, and accuracy of the M1. These improvements have, in turn, expanded the tactical options available to armor unit commanders and individual tank crews. Moving targets may be engaged effectively from a moving M1 tank. The thermal imaging system allows accurate target engagements through smoke, haze, and camouflage as well as during night operations. The M1's laser rangefinder and onboard computer combine to provide the gunner with extremely accurate fire control solutions.

The crews who will man the M1, particularly those with previous armor experience, face a new and challenging armor system. Experienced crews will find that as the technology has changed many of the familiar operating procedures have changed or are no longer appropriate. These crews should be particularly sensitive to the changes. New technologies also represent new theories of operation, new kinds of failures and malfunctions, different diagnostic approaches, and different corrective or compensatory actions when failures do occur. The full benefit of advanced technologies in combat cannot be realized without a capable and well trained crew to operate and coordinate the use of these new systems. The problem is to achieve high levels of M1 crew readiness as quickly as possible with a minimum expenditure of training resources.

Personnel selection is an important factor in attaining combat ready status. Good selection methods can identify those individuals most likely to successfully complete training and to perform effectively in operational units after training. When poor selection decisions are made, training resources are expended where the likelihood of training success is minimal. In the case of tank crews, a poor decision in crew position assignment can delay, if not prevent, the achievement of overall crew readiness. A variety of techniques are available to support selection decisions. These include the use of interviews, application blanks, letters of recommendation, academic records, work records, and personnel testing. The technique which has the greatest potential for avoiding personal bias and inaccurate information and for providing the most reliable and valid information is personnel testing.

Many personnel selection decisions, particularly those in the Armed Forces, are based on an individual's performance on some type of paper and pencil test. The Army uses aptitude and achievement tests to make recruiting decisions, career field assignments, and training assignments. The Armed Forces Qualification Test (AFQT) and the Armed Services Vocational Aptitude Battery (ASVAB) are examples of tests used to support personnel selection decisions. These tests are usually administered at the start of an individual's army career and are used to make initial decisions about the direction that career may take.

Another category of tests, job sample tests, measures "hands-on" performance. Although job sample tests can be designed in paper and pencil formats, they often require performance involving some type of job-related equipment or apparatus. A job sample test is just what the name implies: a sample of the job. Performance on this job sample would, ideally, predict overall job performance.

The Army Research Institute (ARI) has responded to the Army's need to carefully select and assign personnel to armor crews. ARI has investigated both paper and pencil tests (e.g., Eaton, Bessemer, and Krisiansen, 1979) and job sample tests (Eaton, 1978; Eaton, Johnson, and Black, 1980) as predictors of armor gunnery performance. While most of the armor personnel selection research has been performed using paper and pencil tests, job sample tests (Eaton, Johnson, and Black, 1980) have shown promise as predictors of gunnery performance. Guion (1979a, 1979b, 1979c, 1979d) has written extensively on work sample testing and has provided guidelines for the development of work sample or job sample tests. This present research followed Guion's guidelines in the development and validation of a battery of job sample tests to predict M1 gunnery performance.

The specific objectives of this research were divided into two phases. The objectives for Phase I were:

- Develop an aptitude measurement methodology which could be used to design job sample tests for armor crewmen.
- Use the aptitude measurement methodology to develop a battery of job sample tests.
- Administer the job sample tests to experienced M60A1 crews prior to their transition to the M1.
- Collect biographic data and standardized test scores for the M60A1 crews.
- Analyze the job sample test data, the biographic data, and the standardized test scores.

Because of changes to M1 transition training schedules, criteria data for the validation effort were not immediately available. The validation effort was Phase II. Its objectives were:

- Collect M1 transition training performance data for those crews tested in Phase I.
- Collect M1 Table VIII firing data for those crews tested in Phase I.
- Analyze relationships between predictors (Phase I) and performance criteria (Phase II).

The results of the Phase II effort are contained in a separate technical report.

Section 2

DEVELOPMENT OF AN APTITUDE MEASUREMENT METHODOLOGY

One of the objectives of this research effort was to develop an aptitude measurement methodology. Specifically, a methodology was needed to guide the development of job sample tests which could be used as measures of aptitude and predictors of gunnery performance for the M1 tank commander and gunner.

Aptitudes are abilities which develop through heredity and growth and which depend upon the extent to which these have been improved through exercise and experience in general. Aptitude measurement, within the context of personnel selection, represents a forecast of an individual's performance in some future job assignment. An aptitude, then, refers to a person's potential for succeeding in a job prior to training in a specific task (Maier, 1973). In aptitude measurement, measures of an individual's performance prior to training are taken as an indication of how well he will perform after training.

Despite the emphasis upon aptitude measurement, it was felt that focusing upon the concept of "aptitudes" is not a very productive means of developing job sample tests. First, an aptitude is an intervening variable which must be inferred from observed behaviors under given stimulus conditions. Aptitude measurement, then, involves identifying the underlying aptitude, making logical deductions about the behavior which should result under different stimulus conditions, and confirming that the presumed relationships do indeed exist. To say that it is a given aptitude and not some other, the presumed relationships must be mutually exclusive from those predicted by other aptitudes. Often, however, different aptitudes lead to the same deductions about the presumed relationships. Added to this is the problem of aptitude nomenclature--what to title the given aptitude and whose aptitude identification scheme should be employed.

Identification and measurement of a specific aptitude is a problem of construct validity and as such, depends upon the accumulation of evidence which confirms presumed relationships. Concern over specific aptitudes is more constructive <u>after</u> the predictor-criteria relationships have been well defined. In that the aptitude measurement involves an inference from observed behavior and that different aptitudes often lead to the same deductions about the relationships which must exist, it was concluded that identification of specific underlying aptitudes is not a fruitful avenue to job sample development.

Second, concern over whether one is measuring an aptitude places one in a logical bind. To claim that one is measuring solely an aptitude involves showing that either experience is held constant or that differences in experience do not result in differences in performance. This is tantamount to proving the null hypothesis; that is, invoking the concept of aptitude rests upon failure to find significant differences due to experience. However, lack of significance can be produced by other factors such as insensitive metrics, large error variance, and insufficient sample size. If one fails to find significant effects of experience, the conclusion that one is measuring aptitudes is tenable but not conclusively demonstrated. Therefore, emphasis on aptitude measurement is not too conducive to job sample development.

Rather than manifest concern over inferred constructs, it was felt that a different approach was required--one that emphasized observed relationships between behavior and the stimulus conditions under which it occurs. In that tank gunnery performance is basically a man-machine interface problem, the present investigation utilized a human factors approach to job sample development. This approach emphasized analysis of task equipment requirements, task behavioral requirements, and the stimulus conditions required for task performance.

The human factors approach to job sample development is depicted in Figure 2-1. It involves five stages: the task analysis stage; the task prioritization stage; the job sample dimensional analysis stage; the trade-off analysis stage; and detailed test development stage.

In implementing the human factors approach, an attempt was made to incorporate Guion's (1979c, pp. 80-82) seven principles of work (job) sample test development. These principles are summarized as:

1. The choice of job content domain needs to be justified.

2. The test content and job content domain should be as congruent as possible.

3. Scoring procedures should approximate formal, fundamental measurement as much as possible.

4. Levels of proficiency should be measured: scores should not be merely dichotomies.

5. The opportunities for irrelevant influences on individual scores should be at a minimum.

6. Work sample scoring, if the test is to be used in an organization with diverse locations, should be standardized on a scale of reference that is applicable to an organization-wide population.

7. Scores from work sample testing in the usual conditions of institutional control must generalize not only to field settings but to a variety of field settings.

The human factors approach addressed Guidelines 1 through 5 in Phase I of this research. Guidelines 6 and 7 are more appropriate for later stages of job sample development, once the validity of the job samples has been established.

The recommended approach for each stage is based on a retrospective examination of the utility of the steps actually used to develop the job

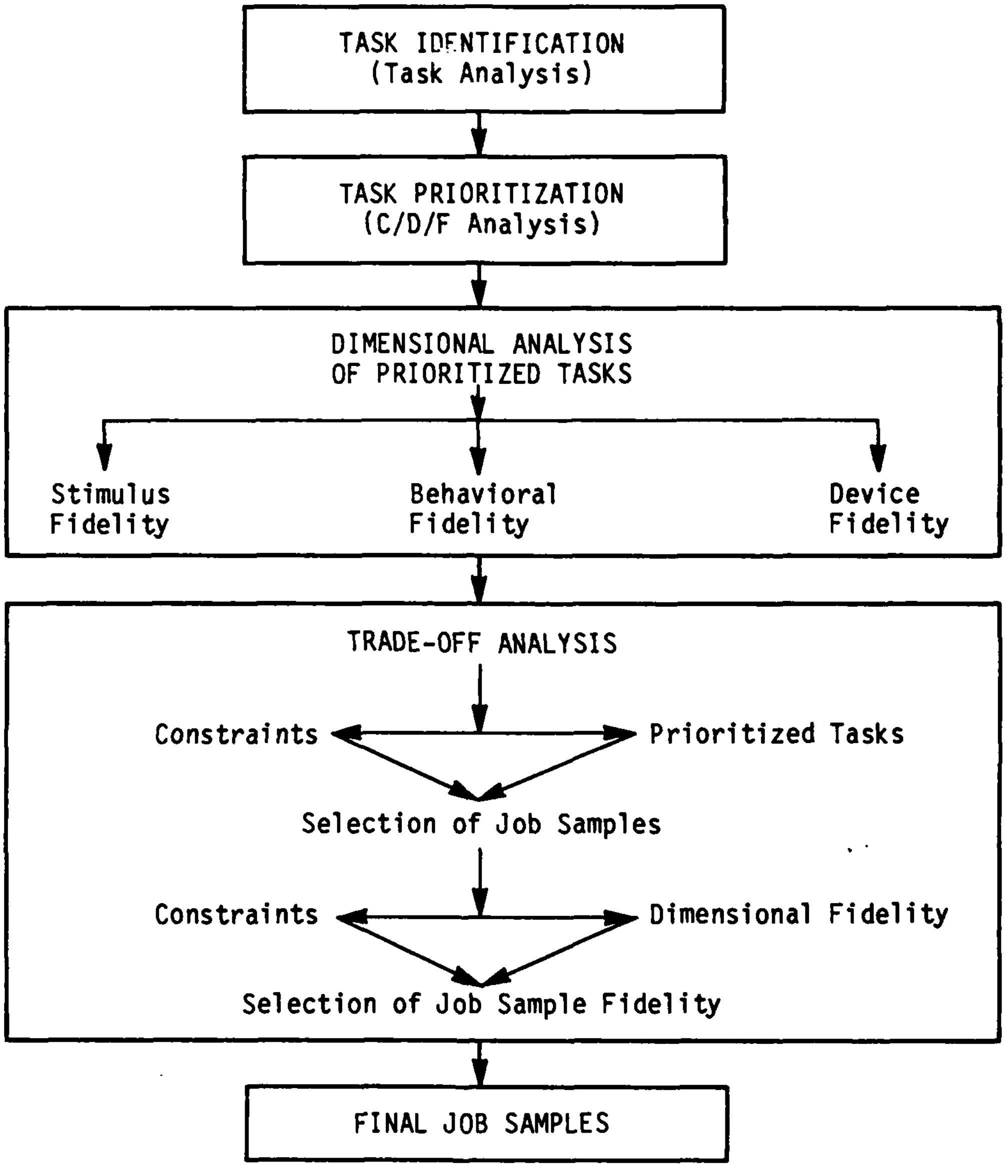

Figure 2-1. Human Factors Approach to Job Sample Development

sample tests in this current research effort. The description of each stage includes the general steps which should be accomplished in that stage. Specific examples will be cited from the current research effort to further illustrate and define procedures.

2.1 TASK IDENTIFICATION

2.1.1 General Approach

The most crucial requirement of a job sample test is that it be representative of the domain of content from which it is drawn. Therefore, as a first step, the domain of content must be identified. This is also the first step in determining and justifying the job content domain, Guion's (1979c) first guideline for job sample development.

The first stage is the task identification stage. It is during this stage that as much task information as possible is collected. Task analyses are used to gather the bulk of this task data. While there is no single task analysis procedure which will be recommended, the task analytic procedure should represent a systematic approach to gathering task information. The procedure should yield information on the equipment or tools used in the task, the individual and crew behaviors required, the stimuli which trigger certain behaviors, and the type of feedback information which indicates degree of successful task performance. It should also identify the types of skills necessary to perform the task.

Some of the methods which may be used include review of technical orders or technical manuals, review of operating manuals and procedures guides, review of job descriptions, observation of job performance, and interviews with job incumbents.

To get the most information about a group of tasks, it is best to use as many of these approaches as possible. Each approach has the potential for adding more task information and generating additional insight into job performance. This will be an especially useful approach if the job sample tests are being developed for new equipment or weapons systems. In this situation, the amount of task information contained in technical manuals or operating procedures is limited. Similarly, fewer operators or crewmembers are available. Those who are available have relatively little experience with the new system.

New systems represent another challenge to the performance of task analyses. In addition to the extra time needed to consult a variety of sources, the sources will often give conflicting information. It is advisable to allow extra time for conducting task analyses for new systems to ensure that as many of the conflicting bits of information as possible can be resolved. A new system probably does not have complete documentation available. System changes are issued in fast succession. Members of a trained crew may not have received equal training, and valuable operating experiences gained in initial test and evaluation may not have been incorporated into existing training programs. Finally, the effect of time on the tasks cannot be fully evaluated. Crews will not have the depth of experience to identify and exploit the subtle idiosyncracies of the equipment. Nor will they have had experience on worn equipment and the changes or modifications to procedures used to compensate for the worn equipment. Conducting task analyses for new systems presents a special set of challenges which must be recognized and included in the task analysis approach and schedule.

2.1.2 Application of Approach to M1 Job Sample Design

In this research effort, the tasks of interest were those involved in M1 tank gunnery. The starting point for the task analyses was the task listing comparing M60 series gunnery tasks with M1 gunnery tasks prepared by Black and Kraemer (1981). This listing served to direct additional analysis efforts which included review of the M1 technical manual, review of M1 gunnery procedures checklists, and several interviews and data reviews with M1 subject matter experts (SMEs). The SMEs were M1 experienced master-gunners, tank commanders, and gunners stationed at Fort Knox, Kentucky.

Because this effort dealt with a new system, several iterations in the data collection process were necessary. Technical manual reviews were required as additional information or changes were published. Interview sessions with SMEs invariably produced information which conflicted with information gathered earlier.

The task analyses resulted in a detailed understanding of the gunnery tasks including the order of task performance, the conditions and contingencies of task performance, crewmember behaviors, and related equipment or hardware operating characteristics. The task duty areas for tank commanders and gunners are presented in Tables 2-1 and 2-2, respectively.

The final collection of this information represents the completion of the first stage. This information serves as the data base for the remaining stages of the methodology and will be used continually during the job sample test development effort.

2.2 TASK PRIORITIZATION

2.2.1 General Approach

Those task duty areas identified in the first stage must now be placed into some type of priority or hierarchical listing. This is also consistent with Guion's (1979c) first guideline for job sample test development. In any group of tasks directed toward the achievement of a specific goal, there are some tasks which are more important than others in achieving that goal. It is this group of tasks which contain the largest number of candidate tasks for job sample tests. The identification of these tasks, therefore, is important in the ultimate selection of job sample tests.

The approach used to identify the most important tasks is based on an approach described by Cream, Eggemeier, and Klein (1978). Tasks are rated on three dimensions: criticality, the degree to which this task is critical to the accomplishment of the overall job or mission; difficulty, the degree

TABLE 2-1. TANK COMMANDER GUNNERY-RELATED TASK DUTY AREAS

Operate Commander's Manual Range Controls
Perform Prepare-to-Fire Checks
Target Engagements with Main Gun (normal) (GPSE)
Target Engagements with Main Gun (normal/emergency) (GPSE)
Target Engagements with Main Gun (emergency) (GSE)
Target Engagements with Main Gun (manual)
Engage Targets Using Battlesight Gunnery (GPSE)
Engage Targets Using Range Data Card
Respond to Specific Fire Control System Failures
Target Engagements with Coax (normal/emergency) (GPSE)
Target Engagements with Coax (normal/emergency) (TIS)
Target Engagements with CAL .50 (normal)
Target Engagements with CAL .50 (manual)
Engage Multiple/Simultaneous Targets
Troubleshoot Turret
Perform During-Fire PMCS
Perform Post-Fire PMCS

TABLE 2-2. GUNNER GUNNERY-RELATED TASK DUTY AREAS

Perform GPS Functional Check
Perform TIS Checkout
Operate Computer Control Panel
Operate the Ballistic Computer
Test Fire Control System
Conduct Master Gunner Detailed Checks
Operate Muzzle Reference System
Operate LASER Rangefinder
Boresight Armament
Zero Armament
Perform Prepare-to-Fire Checks
Target Acquisition
Target Engagements with Main Gun (normal) (GPS)
Target Engagements with Main Gun (normal/emergency) (TIS)
Target Engagements with Main Gun (emergency) (GPS)
Target Engagements with Main Gun (manual) (GAS)
Main Gun Target Engagements Using Battlesight Gunnery
Respond to Specific Failures
Troubleshoot Turret
Perform After-Fire PMCS
Lubricate XM1 according to Lubrication Order (LO)

of difficulty involved in performing the task; and frequency, the frequency of task performance. These raters should be SMEs for the tasks under consideration. The raters perform the criticality/difficulty/frequency (C/D/F) analysis using a three-point scale. A rating of one represents low criticality, little difficulty, or low frequency of performance. A rating of three represents the highly critical, very difficult, and very frequently occurring task.

Each task is rated on the three dimensions, and an overall rating is calculated by adding the ratings on the dimensions. Scores can range from three (a task which is not critical, is easy to perform, and is performed infrequently) to a nine (a highly critical, very difficult task which is performed very frequently). The original use of these types of analyses was to identify those tasks which should be considered when specifying design requirements for simulators. The approach was used here to identify those tasks which were most important for job or mission success and, therefore, represented strong candidates to serve as the basis for job sample tests.

There is no fixed score for determining which tasks should be retained for and which tasks should be dropped from further consideration. Those tasks rated in the upper half or upper third may be continued to the next step of the methodology. The number of tasks selected for further analyses will, of course, depend upon the resources available to develop the job sample tests.

The result of this stage will be a list of tasks ordered along the dimension of importance. This ordering will clearly show where initial efforts should and should not be directed. It should also be emphasized that the ordered list can guide selection of additional tasks if initially selected tasks are found not suitable in later stages of the methodology.

2.2.2 Application Approach to M1 Job Sample Design

In the current research effort, SMEs were used in the C/D/F analysis of the M1 tank gunnery task duty areas identified in the first stage. The SMEs were three tank commanders and four gunners with hands-on experience on the XM1 tank. They were given task listings of gunnery tasks and instructed to rate the criticality, difficulty and frequency of occurrence of each task. Tank commanders rated those tasks listed in Table 2-1, and gunners rated those tasks listed in Table 2-2. These ratings were made in the presence of two researchers who were available to answer questions about the rating process and record the C/D/F responses. The ratings of task difficulty and frequency of occurrence seemed to be readily understood and were made with little difficulty.

The SMEs did have problems rating task criticality. The researchers determined that the SMEs could rate criticality under two basic situations: battlefield conditions and training conditions. Some tasks could be rated extremely critical under battlefield conditions. These same tasks, under training conditions, would not be rated as critical. Based on this observation, the C/D/F rating process was amended by asking SMEs to rate task criticality under both battlefield and training conditions.

The C/D/F ratings for each task duty area were summed to arrive at an overall index of task importance. Those tasks which achieved a high overall rating were retained for further analyses. Those tasks which achieved low ratings were not considered candidates for job sample test development and were dropped from further consideration. The criterion used in this research for inclusion in the prioritized task listing was a C/D/F rating in the upper half of all task ratings. Tables 2-3 and 2-4 present the prioritized task listings for the tank commanders and gunners, respectively.

TABLE 2-3. PRIORITIZED LIST OF TANK COMMANDER TASK DUTY AREAS

1.	Troubleshoot Turret
2.	Respond to Specific Fire Control Failures
3.	Perform During-Fire PMCS
4.	Engage Multiple/Simultaneous Targets
5.	Perform Prepare-to-Fire Checks
6.	Target Engagements with CAL .50 (manual)
7.	Target Engagement with Main Gun (manual)
8.	Perform Post-Fire PMCS
9.	Target Engagements with Main Gun (emergency) (GPSE)

TABLE 2-4. PRIORITIZED LIST OF GUNNER TASK DUTY AREAS

1.	Perform TIS Checkout
2.	Operate the Ballistic Computer
3.	Operate Laser Rangefinder
4.	Boresight Armament
5.	Target Acquisition
6.	Target Engagements with Main Gun (normal) (GPS)
7.	Target Engagements with Main Gun (normal/emergency) (TIS)
8.	Target Engagements with Main Gun (manual) (GAS)
9.	Respond to specific failures
10.	Troubleshoot Turret
11.	Perform After-Fire PMCS

2.3 JOB SAMPLE TEST ANALYSIS

2.3.1 General Approach

The third stage should concentrate on those task duty areas which received high ratings on the C/D/F analysis. These are the tasks which should be strong candidates for use as job sample tests. This stage then represents efforts to identify the possible ways in which job sample tests could be designed based on these tasks.

The analysis in this stage was guided by a three-dimensional framework. The analyses of each task focused on the three dimensions which appeared to be important in designing a job sample test. These dimensions represented the stimuli associated with the task, devices associated with the task, and task behavior.

The three dimensions considered are applicable to tasks which involve operation of or interaction with some type of equipment, device, or weapon system. Obviously, devices are not considered for any tasks in the group which do not require interaction with equipment.

Task stimuli include the information coming from the work environment which is directly related to and necessary for task performance. The first step is to identify the stimuli associated with a task. Although much of the information on task stimuli should be available from the task analysis conducted in the first stage and from interviews with SMEs, a thorough review should now be conducted to ensure that all relevant details of task stimuli are identified. It is important to point out that while some task stimuli may be produced by the equipment actually used in the task, they are stimuli nevertheless and should be considered along with other task stimuli. The equipment related stimuli, however, may also be included in analyses of task devices.

When the task stimuli have been identified, the next step is to consider possible ways of presenting each stimulus in a job sample test. These analyses in this stage should not be bounded by technological feasibility. Rather, analysts should seek to identify and list as many approaches as possible.

Similar analyses should be conducted for task devices. A listing should be created of ways to represent that device in a job sample test based on the task. Again, it is important at this point not to restrict the possible ways of presenting the devices. In this research, one of the tasks involved the use of the M1 computer panel. The actual M1 computer panel was listed among the ways of representing this device in a job sample test even though the availability of an actual M1 tank was only a remote possibility. Other ways of representing the M1 computer included mockups, diagrams, drawings, and computer graphics.

Finally, task behaviors should be reexamined to ensure that all behaviors have been identified. Some types of behaviors, such as decision making, can be incorporated directly into job sample tests. Other types of behavior, such as the control movements required to track a target, may only be accurately incorporated into a job sample test if the actual equipment or a simulator were also used in the job sample test. For job sample tests which do not include actual equipment, job sample behaviors similar to the actual task behaviors need to be identified. A target tracking job sample test performed on an operational tank will incorporate the same behaviors used on the actual task. A target tracking job sample test performed on a computer may require different tracking behaviors.

Test developers should be particularly sensitive to job sample test behaviors which depart from actual task behaviors. Guion's second principle of job sample test development states that the "test content domain and the job content domain should be as congruent as possible" (1979c, p. 80). Job sample test behaviors which are not congruent with task behaviors may, in some cases, no longer be samples of the task.

2.3.2 Application Approach to M1 Job Sample Design

Those task duty areas which received high ratings on the C/D/F analyses were considered in this stage of the methodology. The first consideration was to identify the types of stimuli associated with these tasks and identify the ways in which these stimuli could be presented in a job sample test. The stimuli were of two general types: stimuli associated with the external environment and stimuli associated with the tank and its subsystems. Stimuli associated with the external environment included panel targets and target vehicles, terrain and vegetation, hit locations, weather, lighting conditions, and thermal energy. Stimuli associated with the tank included feedback from turret controls, sights, information displays, switch settings, lighted and mechanical indicators, and the tank commander's fire commands.

The possible ways of presenting external stimuli were identified as follows:

- Range Targets
- Scaled Range Targets
- Actual Target Vehicles
- Film or Video Tape of Targets
- Slides of Targets
- Drawings of Targets
- Snakeboard
- Flash Cards
- Computer Graphics

The possible ways of presenting stimuli associated with the tank were identified as follows:

- Actual M60A1 or M1 Tank and Included Equipment
- Tank Simulator
- Turret Trainer
- Mockups
- Equipment Diagrams/Drawings
- Computer Graphics
- Computer Game Controls
- Live Fire
- Live Fire--Subcaliber
- M55 Laser Device
- Actual Fire Commands

- Tape Recorded Fire Commands

- Written Fire Commands

The equipment and devices associated with the important gunnery tasks were considered next. The equipment used in these tasks was located in the turret at either the TC's station or the gunner's station. The possible ways that the equipment could be represented in job sample tests were identified as follows:

- Actual M60A1 or M1 Tank and Included Equipment

- Tank Simulator

- Turret Trainer

- Mockups

- Equipment Diagrams/Drawings

- Computer Graphics

- Computer Game Controls

- M55 Laser Device

As indicated earlier, there may be some duplication when considering ways to present task stimuli and ways to represent task equipment or devices. In this case many of the possible approaches for presenting tank gunnery stimuli are also appropriate for representing tank gunnery equipment.

The final part of the job sample test analyses was directed at task behaviors. Various combinations of stimuli and equipment were considered to assess the impact on task behaviors. For example, if a target engagement job sample test included an actual M60A1 tank, an M55 laser device, and targets represented by slides, it was judged that the test behaviors deviated only slightly from actual task behaviors. Most of the combinations of stimuli and equipment considered maintained a close match between test behaviors and actual task behaviors.

Stimuli and equipment combinations which included austere mockups or computer graphics or computer game controls tended to require test behaviors which did not correspond closely to actual task behaviors. A tracking job sample test performed using computer graphics and computer game controls differed from the behaviors used in the actual tracking task using the gunner's sight and cadillac controls.

2.4 TRADE-OFF ANALYSES

2.4.1 General Approach

Resource limitations, operating schedules and job sample testing locations are some of the factors which inevitably place constraints and limitations on the design of job sample tests. In this stage of the job sample test development methodology, those tasks which have been identified as candidates for job sample tests are evaluated against these constraints. The analyses in this stage are designed to determine the degree of fidelity which can be included in the job sample tests given the nature of the constraints and limitations in effect.

The term fidelity is used here as it is in discussions of training devices and simulators. Fidelity is the degree to which job sample test stimuli, devices, and behaviors represent actual task stimuli, devices, and behaviors. The analyses, which follow the three job sample test dimensions identified in the job sample test analysis stage, include stimulus fidelity, device fidelity, and behavioral fidelity. In this stage, the possible ways of presenting stimuli or representing devices or including behaviors are examined in light of the project constraints and limitations. Tradeoffs between fidelity and constraints are made to determine the dimensions of a job sample test which can be achieved under the constraints. For example, if one of the constraints is that the job sample test must be administered at Armed Forces Entrance Examination stations, the use of actual target vehicles as test stimuli would be out of the question. In this situation it may not even be possible to use slides, films, or video tapes. One or more of the other possible ways to present test stimuli must be considered.

Devices used in job sample tests must be considered in a similar manner. In the example above, it would probably not be possible to bring the actual device to the test station. A mockup of the device or a diagram may have to be substituted in the job sample test.

Behavioral fidelity must also be considered. It is the dimension which is usually considered after the stimuli and device fidelity are determined. This does not mean it is a dimension of lessor importance. As job sample test stimuli and devices move away from actual stimuli and devices, the job sample test behaviors are more likely to diverge from the actual task behaviors. The degree of divergence under the job sample test conditions indicates the lack of behavioral fidelity. The job sample test behaviors must, therefore, be examined closely and critically to determine if the degree of behavioral divergence is acceptable. If it is not acceptable, then the test designers must reconsider stimuli and device fidelity in light of the behavioral fidelity issues. Behavioral fidelity, therefore, can serve as a check on the extent to which stimulus and device fidelity depart from the actual task. Acceptable behavior fidelity further ensures that test behaviors are congruent with task behaviors (Guion, 1979c, p. 80).

Job sample tests which meet project constraints can be defined based on the analyses in this stage. The combination of the stimulus presentation with the method of device representation which preserves behavioral fidelity defines the basic dimensions of a job sample test for a particular task. The analyses cannot necessarily be accomplished sequentially starting with stimulus fidelity, then device fidelity, and then behavioral fidelity. They should be done simultaneously and the overall analysis may require several iterations before a satisfactory set of stimuli, devices, and behaviors can be identified.

2.4.2 Application of Approach to M1 Job Sample Design

In the present research effort, several constraints governed the job sample test design. The cost constraint dictated a low budget for test equipment development and fabrication. Sophisticated simulators and part task trainers were beyond the cost limitations of this effort. Although the use of a microcomputer may also appear to exceed the cost limitations, a single computer can support a variety of job sample tests. In this effort the cost of a computer was spread over three job sample tests.

The cost constraint was also a factor in the decision not to use live firing or full scale ranges where the ammunition and fuel requirements could restrict the number of times job sample tests could be used. The use of parked tanks, the M55 laser device, and target slides fit well within the cost limitations.

Another constraint was that the job sample tests should use equipment available at the field unit level. Again this favored actual tanks, M55 laser devices, slide projectors, and snakeboards which could be locally produced. It was also recognized that computer-based training devices would, in the future, be included in the training resources at the field unit level. This was another reason for including computer-based job sample tests in this current effort.

Another type of constraint which was considered was the type of criteria data which were available for validating the job sample tests. To ensure the best chance of finding predictive relationships between job sample tests and the criteria, the job sample tests should include those types of tasks included in the performance criteria. The major criteria were to be Table VIII qualification scores. These criteria were more consistent with target engagement activities including target acquisition, target identification, tracking, use of the computer control panel, use of the laser rangefinder, and use of the thermal imaging system (TIS). Troubleshooting tasks, responding to degraded or failure modes, and conducting maintenance checklists were not directly measured in the Table VIII qualifications.

Not all constraints can be identified or stated at the start of a project. Often the analyses in this last stage will uncover additional limitations. In considering stimulus fidelity for several of the M1 tank gunnery tasks, it became apparent that there was a very limited selection of infrared target imagery to serve as stimulus material for tasks involving the M1's TIS. The imagery which was available was not suitable for use in a job sample test. Furthermore, analysis of the TIS tasks revealed that high fidelity imagery was crucial and that lower fidelity stimulus material,

material more abstract than actual infrared imagery, would not be acceptable for use in a TIS job sample test. As a result of the limited imagery, it was decided that a TIS job sample test would not be constructed under the current research effort.

The seven job sample tests which were developed met the cost, equipment, and criteria constraints as well as limitations in the capability to present appropriate stimuli. The tests tap target engagement tasks including the use of the computer panel and laser rangefinder. The TIS tasks, trouble-shooting tasks, degraded mode tasks, and maintenance tasks did not meet the limitations and constraints and were not included for further job sample test development.

2.5 DETAILED TEST DEVELOPMENT

2.5.1 General Approach

The first four stages result in the selection of specific job sample tests based on organizational goals and resources. The detailed development of those job sample tests finally selected can then proceed building on the data developed during the first four stages of the methodology. The remaining test design and development will involve editing or designing stimulus material, defining details of testing and scoring procedures, and procuring and constructing any test apparatus or devices. Guion's (1979c) principles should be used throughout this stage. Scoring procedures and metrics should be as close to fundamental measurement as possible. Levels of task proficiency, as opposed to pass/fail or yes/no dichotomies, should be incorporated into the test design. Procedures, measurement devices, and instructions should be designed to achieve standardization in test administration and avoid irrelevant influences on test scores.

2.5.2 Application of Approach to M1 Job Sample Design

This development effort resulted in the design and development of seven job sample tests which were consistent with Guion's guidelines for job sample test development. In the detailed design of the individual job sample tests, the following factors were considered:

- Test Conditions
- Task Description
- Independent Variables
- Procedures
- Dependent Variables
- Scoring
- Equipment
- Approximate Administration Time

The descriptions of the job sample tests which were designed following the five-stage methodology are contained in Section 3 and Appendix A of this report.

Section 3

JOB SAMPLE TEST DESCRIPTION AND PHASE I DATA COLLECTION METHODOLOGY

3.1 SUBJECTS

All subjects were M60A1 tank crewmembers from four companies of an armor unit in Germany. They were scheduled to begin transition training for the M1 tank within two months of the completion of the job sample testing. Tank commanders (TCs) and gunners participated in all seven job sample tests. Loaders and drivers participated in the three computer-based tests only: operation of the M1 computer control panel, M1 target engagement, and computer tracking. Table 3-1 contains the numbers of crewmembers who participated in the job sample testing.

TABLE 3-1. NUMBER AND TYPES OF SUBJECTS TESTED

Crew Position	Microcomputer-Based Job Sample Tests	Hands-On Job Sample Tests Only
Tank Commander	35	35
Gunner	53	53
Loader	57	0
Driver	44	0
Total	189	88

3.2 TESTS

A total of seven job sample tests were administered. Four of the tests were conducted on the M60A1; and three of the tests were conducted using the Apple II Plus microcomputers, 12-inch color video monitors, a joystick controller, a light pen, a slide projector, and an image combiner box. Each of the seven job sample tests are summarized in this section. Detailed descriptions of the tests, test procedures, and equipment are contained in Appendix A.

3.2.1 <u>Computer-Based Tests (Off-Tank)</u>

3.2.1.1 Computer Panel

Subjects included TCs, gunners, loaders, and drivers. Subjects performed three types of operations on a simulation of the M1 computer control panel. The simulated M1 computer control panel was presented by means of

Apple II Plus computer generated graphics on a 12-inch color video monitor. A light pen was used to operate the simulated control panel. The operation of the computer control panel was demonstrated for each subject. Subjects were then given five practice operations to perform independently. Following the five practice trials, subjects completed 10 trials involving checking data and entering data into the M1 computer. Subjects then completed 10 trials of computer self-test tasks. Performance measures on each trial included whether the trial was performed correctly or incorrectly and the amount of time required to complete each trial.

3.2.1.2 Computer Tracking

Subjects included TCs, gunners, loaders, and drivers. Subjects viewed a 12-inch video monitor which contained a graphic representation of the M1 reticle and a target dot (approximately .16 cm square), which moved in a random pattern across the monitor at one of three target speeds. Subjects used the joystick control to keep the reticle centered on the moving target (reticle was stationary in center of screen). Subjects completed three two-minute trials with the target dot moving at a higher speed on each successive trial. Performance measures were the number of target hits, time on target, and root mean square (RMS) distance error during each trial.

3.2.1.3 Computer Target Engagement

Subjects included TCs, gunners, loaders, and drivers. Subjects performed target location tasks and used M1 target engagement procedures on a computer-controlled, table-top simulation device. An Apple II Plus microcomputer generated graphic simulations of components of the M1 gunner's primary sight, a reticle movable under joystick control, and laser rangefinder information. The computer also controlled the operation of a slide projector containing actual target scenes. An image combiner superimposed the graphic representation of the reticle and laser rangefinder information on the target scene. Subjects were given instructions on the M1 target engagement procedures including the use of the 3X and 10X magnification levels and the use of the laser rangefinder. Instructions concluded with a demonstration of two engagements. Subjects completed 18 trials immediately following the instructions and demonstration. Each trial began with the presentation of a target scene as viewed through the 3X magnification level. Subjects searched the scene for the target and used a joystick to center the reticle on the target. The next step was to change the magnification to 10X by pressing a button on the joystick box labeled 10X. A 10X magnification of the target scene was then presented to the subject. The subject centered the reticle on the target and pressed the button labeled "Laser" to obtain the range to the target. The subject pressed the Laser button a second time if the first Laser button press did not produce valid range data. When valid range data were obtained or after the second Laser button press, the subject pressed the button labeled "Fire" which completed the trial. Performance measures included the distance between the center of the reticle and target location when the 10X, Laser, and Fire buttons were pressed. Time was recorded from the appearance of the 3X target scene to the button presses for 10X, Lase, and Fire. Laser ranging procedures were scored as correct or incorrect for each trial.

3.2.2 Hands-On Tests

3.2.2.1 TC Decision Making

Subjects included TCs and gunners. The subject stood in the TC's hatch and observed a projection screen (6 x 18 feet) placed 18 feet from the end of and perpendicular to the M60A1 main gun barrel. The trial started when three slides of threat vehicles appeared simultaneously on the screen. The subject was instructed to identify the slide which contained the most dangerous threat vehicle. The subject pressed one of three switches corresponding to one of the three slides to indicate his choice of the most dangerous threat vehicle. The trial ended when the choice was made. The screen remained blank while the experimenter recorded the choice of vehicle and the amount of time which elapsed from the simultaneous appearance of the set of three slides to the choice of most dangerous threat. As soon as the choice was made, the slides were advanced to present a blank screen. Subjects were given no information on the accuracy or times of their responses.

3.2.2.2 Hands-On Gun Laying

Subjects included TCs and gunners. Subjects stood in the TC's hatch facing a rectangular screen (5.4m x 1.8m) placed 18 feet from the end of and perpendicular to the M60A1 main gun barrel. A row of three slides was projected simultaneously. Two of the slides were totally clear while the third slide contained a single black dot on a clear background. As soon as the slides appeared, the subject was instructed to use the TC's override control to lay the gun on the black dot. The projected size of the dot was approximately six cm. Special electronic timers were used to measure the time between the appearance of the dot and the completion of the gun laying task as indicated by the release of the TC override palm switch. Accuracy of the gun lay was measured by means of a grid slide and the M55 laser device. When the gun was laid on the dot, the experimenter advanced a grid slide containing a dot in the same position as the target dot on the previous slide. The subject then triggered the M55 laser device to determine the distance between the target dot and the actual gun lay position as indicated by the laser dot. Three practice trials and 12 test trials were presented. The slide containing the target dot appeared an equal number of times at the three slide positions. Within the target slide, the target dot appeared once in each of 12 positions.

3.2.2.3 Hands-On Tracking

Subjects included TCs and gunners. Subjects tracked a snakeboard from both the TC's station, using the TC's override, and the gunner's station, using the gunner's "cadillac" controls and the gunner's sight. A special electronic device was attached to the M55 laser device to cause the laser to flash once every second for a period of one minute. The M55 could not be activated by any of the M60A1 triggers for this test. Subjects began each trial by positioning the steady laser beam at one end of the snakeboard path. The width of the path was 1.8 cm. When the subject was ready, the experimenter switched the M55 to the pulsed mode. Subjects were to track the snake as rapidly and as accurately as possible during the one-minute period when the laser was pulsed once per second. The experimenter measured accuracy by counting the number of laser pulses which appeared (hit) on the snake path. Speed of tracking was measured by the distance traveled on the snake path during the one-minute trial period. Subjects performed six tracking trials from the gunner's station and six tracking trials from the TC's station. The 12 trials were divided into three blocks of four trials. Gunners began a block at the gunner's station, then moved to the TC's station for two trials, and then returned to the gunner's station for the fourth trial. TCs started at the TC's station then moved to the gunner's station for the next two trials and returned to the TC's station for the fourth trial.

3.2.2.4 Hands-On Target Engagement

Subjects included TCs and gunners. Subjects were seated at the gunner's station and viewed a rectangular screen (6 x 18 feet) through the gunner's sight. A qualified TC was at the TC's station as an assistant to the experimenter. A set of three slides was projected simultaneously onto the screen. Two slides contained terrain scenes without targets while the third slide contained a target vehicle on terrain similar to the terrain presented in the other two slides. The TC issued the appropriate fire command and laid the gun on the target. The TC released control of the TC override when the subject called "identify." The subject centered the reticle on target and pulled the trigger, firing the M55 laser device. The TC called hit or miss based on the position of the laser dot relative to the target vehicle. Electronic timers were used to record time from target appearance to release of the TC override and to the firing of the M55 laser device. A total of 15 trials were presented.

3.3 TESTING PROCEDURES

The hands-on job sample tests were conducted over a two-week period at an indoor small-caliber tank firing range. Two M60A1 tanks were parked front to front in the range facility. The main gun tube on each tank was positioned over the tank's rear deck (see Figure 3-1). The hands-on tracking job sample test was conducted using one tank parked 4.5 meters from the snakeboard (see Figure 3-1, Station B). The TC Decision Making, Gun Laying, and Hands-On Target Engagement job sample tests were conducted on the second tank which was parked so that the slide projection screen was 5.4 meters from the end of the gun tube (see Figure 3-1, Station C). All turret operations were conducted using tank battery power or range facility power (24 Vdc). The tank engines were only operated between test sessions to charge the batteries.

During the hands-on testing period, microcomputer-based job sample tests were conducted at two stations in two smaller rooms adjacent to the main room of the range facility. The computer engagement and computer tracking job sample tests were conducted at Station A1 (Figure 3-1). The Computer Panel job sample test was conducted at Station A2 (see Figure 3-1). At the completion of the hands-on testing, the microcomputer-based job sample tests were conducted in a brigade training room.

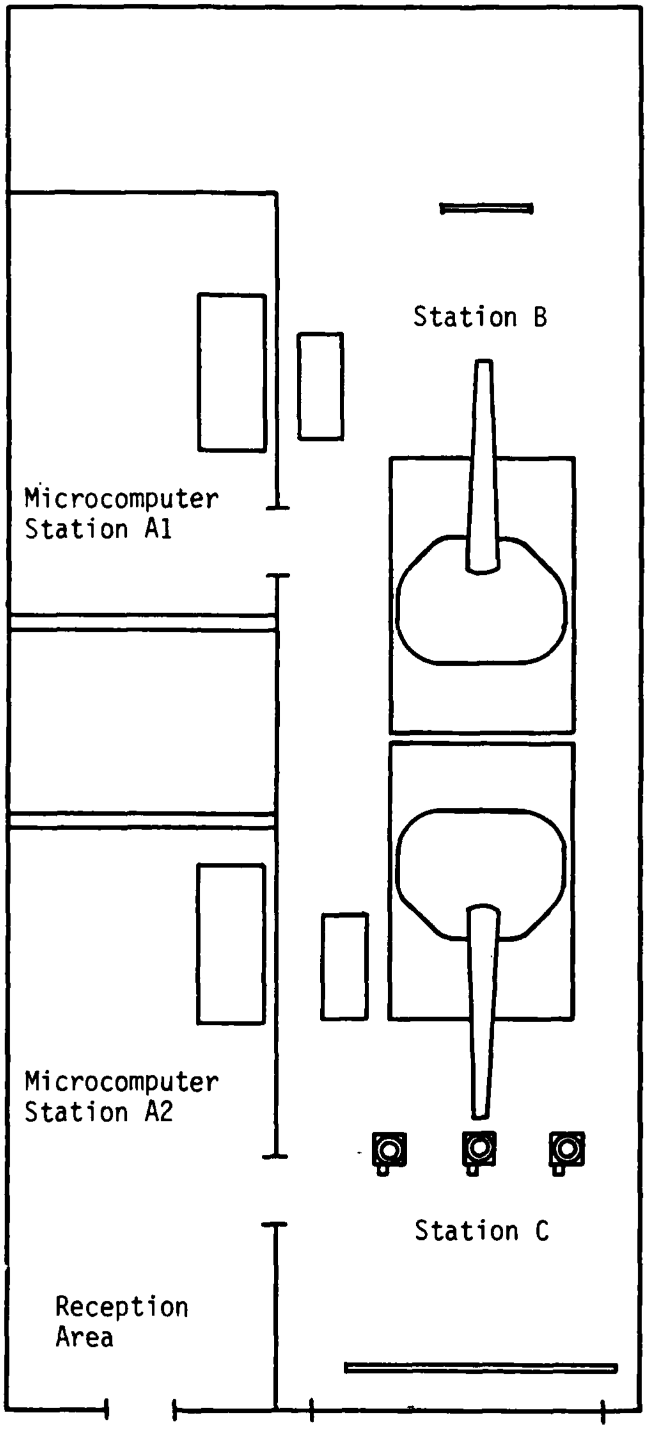

Figure 3-1. Diagram of Indoor Range Facility and Equipment Setup

Subjects reported to the testing location in groups of four at 0730 and at 1300 each day of testing for a four-hour testing session. Several evening test sessions were scheduled to accommodate company operational commitments. The conduct of these sessions was the same as the morning and afternoon sessions. Subjects were scheduled by company so that testing of one company was completed before testing of the next company began.

Upon reporting to the range facility, subjects were assigned a number of from one to four. This number represented the sequence of three testing stations scheduled for that subject. Subjects were rotated in such a way that one subject completed the biographic information form while the remaining three subjects were performing tests at the microcomputer-based test stations, the hands-on tracking test station, or the TC Decision Making/Gun Laying/Hands-On Target Engagement test station.

Only the microcomputer-based tests were administered during the final two weeks of testing. These tests were administered primarily to drivers and loaders. Make-up tests were administered to tank commanders and gunners unable to complete the microcomputer-based tests administered at the range facility. Subjects reported in groups of six on a company basis. Sessions were scheduled at 0730, 1300, and 1800. All subjects completed the biographic data sheet (gunner version) prior to completing the test session. Two subjects immediately began testing at the two testing stations: Operate Computer and Computer Target Engagement/Computer Tracking. The remaining subjects were assigned reporting times during the session at which they were to return for actual testing.

3.4 ADDITIONAL DATA COLLECTION

Additional data was collected from the subject's service records and from their supervisors. Two composite test scores derived from the Armed Services Vocational Aptitude Battery (ASVAB) were collected. These were the combat (CO) score and the general technical (GT) score. In addition, company first sergeants ranked their tank commanders and their gunners on gunnery performance. A ranking of one indicated the best tank commander in the company and the best gunner in the company.

3.5 SCORING

Computer-based job sample tests were scored using data collection routines included in the test program. Scores were recorded on disks and printed out for retention. The hands-on job sample tests were scored by observation and electronic timing devices and were recorded on individual score sheets. Details of the scoring procedures for each test are contained in Appendix A. Score sheet samples are contained in Appendix B.

Section 4

PHASE I: ANALYSIS OF PREDICTOR DATA

4.1 PURPOSE AND SCOPE OF ANALYSIS

The purpose of analysis of the predictor data was threefold:

- To establish the degree of intercorrelation among the individual measures of job sample performance.
- To establish the degree to which experience plays a role in job sample performance.
- To establish the validity of the job samples and the job sample approach in absence of any postpredictor criteria data.

4.1.1 Intercorrelation Among Individual Measures of Job Sample Performance

From a purely tests and measurement perspective, it is highly desirable to have low intercorrelation among predictor measures. Measures which are highly intercorrelated provide redundant (overlapping) information and may tap the same behavior (aptitude). When one uses redundant measures within a prediction framework such as multiple regression, addition of variables to the regression equation is less likely to enhance the ability to predict. On the other hand, if there is a low intercorrelation among the predictor measures, there is a greater probability of finding variables which add to the ability to predict. Thus, with nonredundant measures, it is more likely to find a higher predictor-criterion relationship.

Pearson-product-moment correlation coefficients were employed to determine the interrelationships among the predictors. The criteria of statistical significance ($p = .05$) and the proportion of variability shared in common (r^2) were used to assess the redundancy among the predictor measures.

4.1.2 Role of Experience in Job Sample Performance

The job samples were developed under the presumption that they measure some underlying aptitude. To measure an aptitude, experience and training must be held constant. Therefore, it becomes important to identify the role of experience and training in job sample performance.

The analyses of the role of experience and training proceeded at two levels. First, the role of experience was examined at a gross level group. It may be conjectured that if experience is a factor, groups of individuals having more experience will perform at a different level on a task than groups which do not have that experience. Clearly, drivers and loaders have had little experience on gunnery tasks, and thereby should perform at a different level than either tank commanders or gunners. However, predicting differences in performance of tank commanders and gunners on the basis of experience is not clear cut and is contingent upon the particular job sample. For the computer panel job sample, experience is held constant because it represents a new gunnery task. On the other hand, tank commanders are clearly more practiced and should do better at TC decision making, gun laying, and tracking from the TC station. For hands-on tracking at the gunner's station and hands-on target engagement, the tank commanders are obviously more experienced (in terms of length of service), but this may be counteracted by the recency of their practice at the gunner's station. For the remaining two job samples, Computer Tracking and Computer Target Engagement, it is difficult to generate prediction based upon experience since these are new tasks in that they represent abstractions of actual hands-on tasks.

To assess the role of experience at the group level, two different statistical analyses were used. T-tests were employed to examine differences between means in comparisons involving only tank commanders and gunners. In comparisons involving all four tank crewmen, analyses of variance was used to evaluate differences in group means. It must be mentioned that, in several cases, there was significant heterogeneity of variance. In those instances, appropriate transformations were applied and the analyses redone. In no case did the statistical conclusion about significance change when using the transformed data. Therefore, only the analyses performed on the untransformed data are reported here.

The second level at which the role of experience and training was assessed involved correlating (Pearson r's) numerical measures of the length of past experience (e.g., time in position, time in M60A1 tanks, time in Army, rank) and the recency of training with measures of job sample performance. This correlational analysis was conducted separately for tank commanders and gunners, thus removing any gross group differences in experience and training. Furthermore, evaluation of the role of experience within samples is more appropriate than evaluation between samples because ultimately the predictor-criterion relationships will be derived separately for tank commanders and gunners.

Despite the appropriateness of computing correlations separately for tank commanders and gunners, correlations for the combined subgroups were computed as well. The reasoning behind this is as follows. When looking at the separate correlations for tank commanders and gunners, one may fail to find significant differences because of the lack of power resulting from smaller sample sizes. However, if tank commanders and gunners exhibit similar but nonsignificant correlations, a more powerful test can be made by statistically assessing the significance of the correlation for the combined subgroups.

Interpreting the results of these analyses regarding experience and training is difficult. One wants to conclude that aptitudes and not experience or training is being measured by these job samples. The job samples may be affected by experience or training; but as long as experience and training are held constant, the notion that one is dealing with aptitudes can at least be entertained. However, there is a logical flaw in this analysis--to entertain the hypothesis that one is tapping aptitudes and not experience is tantamount to proving the null hypothesis. That is, invoking the concept of aptitude rests upon failure to find significant differences due to experience and training. Failure to find significant differences leads the unwary observer to the deduction that either experience and training have no effect or that experience and training are equivalent between groups. But the lack of significance could also be produced by insufficient sample size, large error variance, lack of variation in one of the measures, or insensitive measures. Therefore, if one fails to find significant effect of experience and training, the conjecture that one is dealing with aptitude measurement is tenable but not conclusively demonstrated.

On the other side of the coin, if performance on the job samples is a function of experience and training, one can reject the hypothesis that the job samples are solely measuring aptitudes.

4.1.3 Validity of the Job Samples and Job Sample Approach in Absence of Postpredictor Criteria Data

The job samples have immediate content validity due to the manner in which the job samples were selected; namely, the job samples were chosen so that they would be representative of the domain of content of tank gunnery skills and behaviors. In Phase I data analysis, an attempt was made to assess the other two major types of validity as well--construct validity and criterion-related validity.

Construct validity is, in part, a matter of making logical deductions about the relationship between two variables based upon some underlying construct and confirming that that relationship does indeed exist; that is, given knowledge of the underlying construct certain behaviors should be related in a predictable way. If they are not, serious doubt is cast upon the validity of the construct.

In assessing construct validity, interrelationships among predictor measures were examined as to their plausability. At the level of individual predictor measures, bivariate (zero-order) correlations were evaluated as to their directionality and significance.

In addition to individual job sample measures being related in a predictable way, certain job samples as a whole should be related to one another. For example, performance as a whole on the Computer Tracking and Hands-On Tracking job samples should be correlated because they were developed to measure the same gunnery skills at two different levels of abstracton. The same can be said for the Computer Target Engagement and the Hands-On Target Engagement Job Samples. To test the relationship between job samples as a whole, canonical correlations were computed. Canonical correlation is a statistic which measures the relationship between two sets of variables rather than pairs of individual variables as is the case with bivariate correlation. In canonical correlation, a set of weights is derived which maximizes the correlation between the two sets of variables.

The canonical correlations were computed merely to make a decision as to whether one can conclude the set of measures of one job sample are related to the set of measures of another job sample. For purposes of Phase I data analysis, no interpretative use was made of the canonical weights or the correlations of individual variables with the canonical composites.

The construct validity of the job samples can be enhanced if they relate to experience and training in predictable ways. That is, groups with greater experience and training (tank commanders and gunners) should perform at higher levels than groups (drivers and loaders) with less experience. To assess the role of experience and training, zero-order and canonical corelations were computed between measures of experience and training and measures of job sample performance.

At a macro level, one could claim that a single construct underlies all job samples--namely, tank gunnery skill. This is not unreasonable in that tank gunnery, as well as each job sample, involves an emphasis on speed and accuracy. Therefore, to establish the construct validity of the job samples, one should find a high intercorrelation among job samples both at the level of individual predictors (zero-order correlation) and at the level of sets of variables (canonical correlation). However, failure to find significant correlations among individual measures does not invalidate the job samples. It may be that the job samples measure different constructs (i.e., aptitudes).

In that each job sample places an emphasis on speed and accuracy, there should be some commonality among measures of job sample performance; that is, they should be correlated. Again, there is a logical contradiction--what one desires from a test and measurement perspective (uncorrelated measures) differs from what one desires from a validity perspective (correlated measures). Despite this apparent contradiction, it is possible that the results could end up supporting both notions. For example, under the assumption that different job samples involve different constructs, one would expect to find higher intercorrelations among measures within the same job sample (enhancing construct validity) and relatively low intercorrelation across different job samples (good from the testing perspective). On the other hand, it may be possible that there is a low level of intercorrelation among individual measures (bivariate correlation) but a high level of intercorrelation among sets of measures (canonical correlation); that is, when simultaneously considering sets of variables (canonical correlation), a pattern among the sets of variables emerges which was not detected statistically at the level of individual measures. It is the pattern which is correlated, not the individual measures.

In absence of any postpredictor criteria data, criterion-related validity was assessed by working backward in time. The job samples were related to measures of past success at annual qualifications. This analysis proceeded at two levels--by using bivariate correlaton (zero-order) to relate individual measures of job sample performance to the appropriate measure of past success and by using multiple correlation to relate a set of job sample measures to the appropriate measure of past success. The multiple correlations were performed on a job sample by job sample basis. Measures of experience and training were also correlated (both bivariate and multiple correlation) with past success at annual qualifications to establish whether the job samples or experience and training accounted for more variability in the measures of past success. If the job samples are more highly correlated with past success than either experience or training, the utility of the job sample approach is somewhat validated.

Finally, to further establish the utility of the job sample approach, a multiple regression analysis was performed regressing past success at annual qualifications on the measures of job sample performance. This regression analysis considered the major measures from all job samples as candidates for entry in a stepwise fashion.

4.1.4 Organization of Analyses and Relationship to Stated Purposes

The presentation of the results is organized around four sets of analyses. The first set of analyses involved examination of the biographical data to determine the extent to which tank commanders and gunners differ on measures of experience and past success at Annual Qualification. In the second set of analyses, each individual job sample was examined to establish the degree of correlation and redundancy among measures within the job sample, to determine the plausibility of relationships among job sample measures (construct validity), and to establish the role of experience in job sample performance. The third set of analyses investigated the interrelationship among different job samples to establish the degree of intercorrelation among job sample measures and to determine the plausibility of the relationships which emerged (construct validity). The final set of analyses involved examination of the relationship of job sample measures with past success at Annual Qualification to determine the criterion-related validity of the job samples.

4.2 ANALYSIS OF BIOGRAPHICAL DATA

The biographical data was analyzed primarily to determine the extent to which tank commanders and gunners differ on measures of past experience. Table 4-1 presents a list of the dependent measures which were derived from the biographical data sheet along with the names used to describe the variables and, where appropriate, the codes assigned to the levels of the variable. To facilitate exposition, these 18 measures are grouped into the following categories:

- General Information (AGE, EDUC)
- Level of Army Experience (RANK, ARMY:TIME, A1:TIME, A3:TIME, CP:TIME)
- Recency of Training (SC:MLAST, VRT:MLAST, CTT:MLAST)
- ASVAB Scores (CO, GT)

TABLE 4-1. DEPENDENT MEASURES DERIVED FROM BIOGRAPHICAL DATA SHEET

Variable Name	Description	Code Levels
AGE	Age	
EDUC	Highest Level of Education Attained	1 = Attended High School 2 = High School Graduate 3 = Attended College 4 = College Graduate
RANK	Rank (Pay Grade)	1 = E1 2 = E2 3 = E3 4 = E4, SP4 5 = E5, SP5 6 = E6 7 = E7
ARMY:TIME	Number Months in Army	
A1:TIME	Number Months Served in M60A1	
A3:TIME	Number Months Served in M60A3	
CP:TIME	Number Months in Current Crew Position	
SC:MLAST	Number Months Since Last Subcaliber Fire	
VRT:MLAST	Number Months Since Last Vehicle Recognition Training	
CTT:MLAST	Number Months Since Last Combat Training Theater (CTT) Training	
CO	Combat Composite Score from ASVAB	
GT	General Technical Composite Score from ASVAB	

TABLE 4-1. DEPENDENT MEASURES DERIVED FROM BIOGRAPHICAL DATA SHEET (continued)

Variable Name	Description	Code Levels
GAME:FREQ	Frequency With Which Play Computer Games	1 = Once Per Month 2 = Once Per Week 3 = More Than Once a Week 4 = Every Day
QAVG:TC	Average Score at Annual Qualifications During 1974-1981 when in Tank Commander Crew Position	1 = Unqualified 2 = Qualified 3 = Distinguished
QAVG:G	Average Score at Annual Qualifications During 1974-1981 when in Gunner Crew Position	See QAVG:TC
QAVG:TCG	Average Score at Annual Qualifications During 1974-1981 when in Either Tank Commander or Gunner Crew Position	See QAVG:TC
MRQ:TC	Score at Most Recent (1981) Annual Qualification when in TC Crew Position	See QAVG:TC
MRQ:G	Score at Most Recent (1981) Annual Qualification when in Gunner Crew Position	See QAVG:TC
MRQ:TCG	Score at Most Recent (1981) Annual Qualification when in Either Tank Commander or Gunner Crew Position	See QAVG:TC

- Computer Games (GAME:FREQ)

- Past Success at Annual Qualification (QAVG:TC, QAVG:G, QAVG:TCG, MRQ:TC, MRQ:G, MRQ:TCG)

The means and standard deviations for the tank commanders and gunners on the 18 biographical measures are shown in Table 4-2. The bivariate correlations (zero-order) among the biographical measures can be found in Appendix C, Table C-4.

As can be seen in Table 4-2, the tank commanders are older than the gunners, but do not differ from them in educational level. The average level of education for both tank commanders and gunners is that of a high school graduate.

As would be expected, the tank commanders have significantly more army experience than the gunners. The average rank for tank commanders is between an E-5 and and E-6 whereas that for the gunners is between an E-4 and E-5. The tank commanders have spent more time in the army, more time in the M60A1, and more time in their current crew position than have the gunners. Tank commanders and gunners do not differ, however, in time served in the M60A3.

It is also important to note that there are significant differences in the range of experience for tank commanders and gunners on each of the five experience measures. This heterogeneity of variance will affect the magnitude of job sample correlations with experience. More specifically, differential correlations with experience for tank commanders and gunners may be obtained simply because of differential variability in these measures.

The experience measures are correlated in accordance with expectation. RANK, ARMY:TIME, A1:TIME, and CP:TIME are all positively intercorrelated and each is negatively correlated with A3:TIME. The longer the time in the Army, the higher the rank, the more time spent in the current crew position, and the more time spent in the M60A1. However as indexed by RANK, ARMY:TIME, and A1:TIME, the greater the length of service, the less time spent in the M60A3.

Although gunners have more recently participated in subcaliber fire, vehicle recognition training, and CTT training, there are no significant differences between tank commanders and gunners in recency of training. Interestingly, only 21 of the tank commanders (n=35) and only 16 of the gunners (n=53) reported they had any recent practice on the CTT device, a trainer whose device and stimulus fidelity somewhat resembles that of the Hands-On Target Engagement job sample.

Since tank gunnery tasks have elements in common with those of computer games, the tank crewmen were asked the frequency with which they play computer games. The presumption is that playing computer games is a form of practice; therefore, the more frequently one plays computer games, the better they will perform in gunnery tasks. Tank commanders and gunners do not significantly differ in the reported frequency with which they play computer games, the average being once a week.

TABLE 4-2. BIOGRAPHICAL DATA: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 18 DEPENDENT MEASURES

Dependent Measures		TC (n=35)	G (n=52)	F^1 Value	t^2 Value
	n	35	52		
AGE	M	28.23	23.38		6.08***
	SD	4.09	3.32	1.52	
	n	35	52		
EDUC	M	2.03	1.87		1.91
	SD	0.38	0.40	1.08	
	n	35	52		
RANK	M	5.86	4.42		11.58***
	SD	0.43	0.72	2.83**	
	n	34	52		
ARMY:TIME	M	85.03	46.19		4.49***
	SD	45.89	25.97	3.12***	
	n	35	53		
A1:TIME	M	61.94	29.92		5.26***
	SD	32.56	18.73	3.02***	
	n	35	52		
A3:TIME	M	0.97	0.73		0.40
	SD	2.23	3.44	2.38**	
	n	35	52		
CP:TIME	M	41.14	15.21		4.89***
	SD	28.44	16.12	3.11***	
	n	32	42		
SC:MLAST	M	15.03	10.45		1.46
	SD	15.19	11.85	1.64	
	n	32	47		
VRT:MLAST	M	6.22	5.53		0.54
	SD	5.85	5.41	1.17	
	n	21	16		
CTT:MLAST	M	11.62	8.13		1.07
	SD	13.12	6.16	4.53**	
	n	29	44		
CO	M	101.76	109.91		-1.93
	SD	15.15	19.12	1.59	

TABLE 4-2. BIOGRAPHICAL DATA: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 18 DEPENDENT MEASURES (continued)

Dependent Measures		TC (n=35)	G (n=52)	F[1] Value	t[2] Value
GT	n	33	43		
	M	109.91	104.23		1.30
	SD	21.55	14.71	2.15*	
GAME:FREQ	n	34	51		
	M	2.24	2.00		1.04
	SD	1.04	1.00	1.09	
QAVG:TC	n	32	7		
	M	2.36	2.04		1.79
	SD	0.44	0.47	1.14	
QAVG:G	n				
	M	2.29	2.22		0.51
	SD	0.54	0.36	2.25*	
QAVG:TCG	n	35	35		
	M	2.28	2.16		1.47
	SD	0.36	0.33	1.17	
MRQ:TC	n	24	6		
	M	2.33	2.00		1.43
	SD	0.48	0.63	1.72	
MRQ:G	n	3	21		
	M	2.00	2.05		0.37
	SD	0.00	0.22	0.00	
MRQ:TCG	n	27	27		
	M	2.30	2.04		2.34*
	SD	0.47	0.34	1.90	

* $p \leq .05$
** $p \leq .01$
*** $p \leq .001$

[1] Tests significance of homogeneity of variance assumption.

[2] Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

The CO and GT scores of the Armed Services Vocational Aptitude Battery (ASVAB) have been frequently used as predictors of gunnery success. In the present study, the tank commanders and gunners do not significantly differ on the average in either composite score, although the tank commanders are significantly more variable in their GT composite scores.

Of the six measures of past success at Annual Qualifications, only one of them shows a significant difference between tank commanders and gunners--MRQ:TCG. In the most recent (1981) Annual Qualification, the tank commanders were members of crews which did better than did crews of which gunners were members. Since most of the tank commanders were at the TC station (24 of 27) and most of the gunners were at the gunner station (21 of 27), the relationship may be taken as representing performance of the tank commanders at the TC station versus performance of the gunners at the gunner station.

It is interesting to note the sample sizes upon which the qualification scores are based. As Table 4-2 indicates, eight of the tank commanders and 26 of the gunners reported they did not participate in the 1981 Annual Qualification as either a tank commander or a gunner. If these crewmen correctly reported their qualification record, a surprising number of tank commanders and a large number of gunners never had experience under pressure in the task which represents the major criterion variable to be analyzed in Phase II. Furthermore, six crewmen who had been to Annual Qualifications as a tank commander in 1981 were now assigned to the gunner position. This suggests a potential motivational problem.

4.3 ANALYSIS OF INDIVIDUAL JOB SAMPLES

The focus in the analysis of individual job samples was threefold:

- To establish the degree of correlation and redundancy among the measures within the job sample.
- To determine the plausibility of the relationships which emerged among the job sample measures.
- To ascertain the role of experience in job sample performance.

The following three sets of information were used to examine the relationships within each job sample:

- Intercorrelations (zero-order) between the measures of performance for the given job sample.
- Differences between tank commanders, gunners, and drivers/loaders (in the case of the computer job samples) in average level and variability of their performance on the job sample measures.
- Correlations between the measures derived from the biographical data (level of Army experience, recency of

training, ASVAB scores, and computer game frequency) and measures of performance in the given job samples.

The analysis of the interrcorrelations among measures of job sample performance addressed the first two purposes whereas the latter two sets of analyses were directed toward determining the role of experience in job sample performance.

Appendix C contains the correlation matrices interrelating the measures of performance within each job sample as well as the correlation matrices relating the biographical and job sample measures. Results of comparisons among means and variances for tank commanders and gunners are presented in the body of the text, whereas results of comparisons involving drivers and loaders appear in Appendix C.

4.3.1 Computer Panel

Seventy-five of the 88 tank crewmen who were sampled completed the Computer Panel (CP) job sample--27 tank commanders and 38 gunners. Due to scheduling problems, eight tank commanders and 14 gunners did not participate.

Table 4-3 presents a list of the six dependent measures which were derived from the Computer Panel (CP) job sample along with their variable name. A more complete description of these measures can be found in Appendix A. Since the CP job sample was divided into two separate tasks (Enter/Check Data and Self-Test), there is a measure of the number correct and completion time for each task. In addition, for purposes of possibly reducing the size of the variable set, the average number correct and average completion time were also computed.

TABLE 4-3. DEPENDENT MEASURES DERIVED FROM COMPUTER PANEL JOB SAMPLE

Variable Name	Brief Description
ECD:CORR	Number Correct on Enter/Check Data (Maximum = 10)
ECD:TIME	Average Time (seconds) to Complete Enter/Check Data Trial
CST:CORR	Number Correct on Self-Test (maximum = 10)
CST:TIME	Average Time (seconds) to Complete the Self-Test Trial
AVG:CORR	Number Correct Averaged Across Two Tasks
AVG:TIME	Completion Time Averaged Across Two Tasks

It is reasonable to expect that the number correct on the two tasks would be postively correlated, as would the completion time on the two tasks. Furthermore, in a task which does not involve an emphasis on speed, it would be expected that completion time would be negatively correlated with the number correct--namely, those individuals taking longer to complete a trial are less likely to make errors. Although Table C-2 reveals that the direction of the correlations for both tank commanders and gunners are in accordance with these predictions, the correlations are not all significant and in some cases almost zero. Number correct and completion time are significantly correlated on the Self-Test but not on the Enter/Check Data task. The correlations between the number correct and completion time on the Enter/Check Data task are low and almost nonexistent. This difference between the two tasks is indicative that they do indeed represent different tasks; therefore, their measures should not be averaged.

The two time measures are significantly and positively correlated for both tank commanders and gunners. However, when comparing the number correct on the two tasks, there was a significant positive correlation only for tank commanders. The failure to find a significant correlation between the two measures of correctness for gunners again suggests that it is inappropriate to average the measures, particularly in the case of the number correct.

Table 4-4 presents the means and standard deviations for the tank commanders and gunners on the six measures of performance on the CP job sample. Given that the CP job sample represents a task which is new to both tank commanders and gunners, there is no reason to expect that they would differ on any of the measures. Although the gunners have slightly more correct and take slightly less time to complete the two tasks, there is no significant difference between the means for tank commanders and gunners on any of the measures.

The only significant difference between tank commanders and gunners is in terms of the variability of their performance. The tank commanders are significantly more variable than gunners in the number correct on the Enter/Check Data task. The fact that the gunners are less variable on this measure could, in part, be responsible for the failure to find a significant correlation between the two measures of correctness; that is, if one restricts the range of one of the variables involved in a correlation, the magnitude of the correlation is reduced.

On the CP job sample, data were also collected on drivers and loaders. Since differences between drivers and loaders were of no interest, the data for these groups were combined for purposes of statistical analysis. When the drivers/loaders are added to the comparison of tank commanders and gunners (cf, Table C-3), a significant difference emerges on only the ECD:CORR measure. The driver/loaders have fewer correct in the Enter/Check Data task than do either the tank commanders or gunners. The failure to find a significant effect in any of the other measures when drivers/loaders was added to the comparison was not unanticipated in that experience should not necessarily be a factor in a task which is new.

TABLE 4-4. COMPUTER PANEL JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 6 DEPENDENT MEASURES

Dependent Measures		TC (n=27)	G (n=38)	F^1 Value	t^2 Value
ECD:CORR	M	7.70	8.05		-0.68
	SD	2.37	1.51	2.47*	
ECD:TIME	M	37.62	36.21		0.67
	SD	7.37	8.88	1.45	
CST:CORR	M	8.04	8.39		-0.68
	SD	2.26	1.94	1.36	
CST:TIME	M	31.74	30.14		0.66
	SD	10.00	9.52	1.10	
AVG:CORR	M	7.87	8.22		-0.81
	SD	2.00	1.26	2.53**	
AVG:TIME	M	34.68	33.18		0.79
	SD	7.25	7.75	1.14	

* $p \leq .05$
** $p \leq .01$

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

Inspection of Table C-4 (Appendix C) confirms that experience is, in general, unrelated to the measures of performance on the CP job sample. In only two cases, one for tank commanders and one for gunners, did the correlation between measures of experience (RANK, ARMY:TIME, A1:TIME, A3:TIME), and measures of CP performance reach significance. However, the direction of these two significant correlations are opposite to expectation for tank commanders: the longer the time spent in the M60A3, the fewer the number correct on the Enter/Check Data task. For gunners, the higher the rank, the less the average number of correct responses. Not only are these two correlations unanticipated but the other correlations with experience are low and, in many cases, in opposite directions for tank commanders and gunners.

The same pattern emerges when one compares recency of training with the CP performance measures. In general, the correlations are low and, in many cases, opposite in direction for tank commanders and gunners. There is only one significant correlation, and it is opposite in direction to what one might predict. Namely, for tank commanders, the longer since last subcaliber fire, the shorter the completion time on the Enter/Check Data task.

There is one consistent pattern which emerges with a "training" measure--that between GAME:FREQ and the time measures. For both tank commanders and gunners, the more frequently computer games are played, the shorter the completion time. This relationship is significant, however, only for average time.

Upon inspection of the ASVAB test scores, a consistent and interpretable pattern arises for CO scores but not for GT scores. The higher the CO score, the greater the number correct and the shorter the completion time. This relationship is significant, however, only for the number correct measures on the Enter/Check Data task. For the GT scores, there is a consistent, but nonsignificant, negative relationship with the completion time measures. On the other hand, the GT scores and the number correct are not always related in the same fashion for tank commanders and gunners.

In summary, the measures of performance on the CP job sample are related in a predictable fashion. However, given the pattern of intercorrelations, the measures of correctness should not be averaged. Therefore, in presentation of the results in subsequent sections, emphasis is placed on ECD:CORR, ECD:TIME, EST:CORR, and CST:TIME. Finally, experience and recency of training do not appear to be factors in CP job performance.

4.3.2 Computer Tracking

Eighty-three tank crewmen particpated in the Computer Tracking (CT) job sample--32 tank commanders and 51 gunners. However, due to disk storage problems, data was lost for a number of subjects on the Moderate and Hard tracking tasks. For the Moderate tracking task, only one crewman's data is missing, that of a tank commander. The data loss for the Hard tracking task is more severe, with the final sample size being based on 26 tank commanders and 43 gunners.

Eight dependent measures were derived from the CT job sample. These measures are listed in Table 4-5 (see Appendix A for a complete description of the measures). For each of the three CT tasks (Easy, Moderate, Hard), there are two measures-total time-on-target and magnitude of tracking error. Additionally, to evaluate the possibility of reducing the size of the variable set, the average time on target and the average error were calculated.

TABLE 4-5. DEPENDENT MEASURES DERIVED FROM THE COMPUTER TRACKING JOB SAMPLE

Variable Name	Brief Description
EASY:TOT	Time on Target (sec) for Easy Tracking Task
EASY:ERROR	RMS error (number pixels) for Easy Tracking Task
MOD:TOT	Time on Target (sec) for Moderate Tracking Task
MOD:ERROR	RMS Error (number pixels) for Moderate Tracking Task
HARD:TOT	Time on Target (sec) for Hard Tracking Task
HARD:ERROR	RMS Error (number pixels) for Hard Tracking Task
AVG:TOT	Average Time on Target (sec)
AVG:ERROR	Average RMS Error (number pixels)

Inspection of Table C-5 (Appendix C) reveals that all intercorrelations among measures of CT performance are significant with one exception. The correlations are all high, similar for both tank commanders and gunners, and in the direction one might predict. Due to the nature of the measures, time-on-target and tracking error are negatively correlated. Namely, as time-on-target increases, tracking error must necessarily decrease. More importantly, time-on-target is positively correlated across CT tasks, as is tracking error. The high intercorrelation of the measures across tasks suggests it may be possible to reduce the size of the variable set by employing average error and average time-on-target in subsequent regression analyses.

When comparing tank commanders, gunners, and driver/loaders in terms of the level of their performance on the dependent measures, a surprising finding emerges (cf, Table C-6). Tank commanders and gunners are clearly more experienced than either drivers or loaders in tracking. Yet, the group of drivers and loaders spend more time on target and have less tracking error than either tank commanders or gunners. As indicated by Table C-6, there are significant differences among the three groups for all variables except EASY:ERROR and MOD:ERROR.

These significant differences, however, cannot be attributed to differences among tank commanders and gunners. Table 4-6 presents the results of the analyses comparing means and standard deviations of the tank commanders and gunners. Despite the fact that gunners spend more time on target and have less tracking error (with one exception), tank commanders and gunners do

TABLE 4-6. COMPUTER TRACKING JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 8 DEPENDENT MEASURES

Dependent Measures		TC (n=32)	G (n=51)	F[1] Value	t[2] Value
	n	32	51		
EASY:TOT	M	17.56	20.77		-1.24
	SD	12.00	11.05	1.18	
	n	32	51		
EASY:ERROR	M	38.40	39.10		-0.25
	SD	12.56	11.99	1.10	
	n	31	51		
MOD:TOT	M	13.94	14.86		-0.53
	SD	7.65	7.64	1.00	
	n	31	51		
MOD:ERROR	M	29.38	28.98		0.14
	SD	11.26	12.78	1.29	
	n	26	43		
HARD:TOT	M	6.50	7.88		-1.25
	SD	4.75	4.25	1.25	
	n	26	43		
HARD:ERROR	M	30.46	25.64		1.47
	SD	14.91	12.06	1.53	
	n	26	43		
AVG:TOT	M	12.82	14.81		-1.16
	SD	7.34	6.71	1.20	
	n	26	43		
AVG:ERROR	M	33.86	31.30		0.94
	SD	10.52	11.17	1.13	

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

not differ significantly on any of the measures in terms of either their average level of performance or their variability. This is not unexpected in that tank commanders and gunners both have had experience tracking.

The group with less experience in tracking (drivers/loaders) outperformed the more experienced tank commanders and gunners. This suggests some sort of negative transfer problem where previous experience interferes with performance. In that the computer tracking task differs in many important ways from a tracking task on the tank, it would not be unreasonable to expect a negative transfer problem. For example, when the tank commander and gunner track in a tank, there is a visual scene which moves in a direction opposite to the target. However, the CT job sample has no background visual scene. Furthermore, the computer joystick controls are unlike those of both the tank commander's power control handle and the gunner's cadillac control.

That there may be a negative transfer problem is supported by the correlation of the measures of experience with CT performance (cf, Table C-7). The measures of experience are, in general, negatively correlated with time on target and positively correlated with tracking errors. These relationships are strongest for time spent on the M60A1, time spent in current crew position, and for tank commanders. Significance was obtained primarily on A1:TIME and CP:TIME measures.

The picture regarding recency of training does not fit any predictable pattern. One would expect, based upon the results of the experience measures (i.e., negative transfer), that the longer the time since last training exercise, the more time on target (positive correlation) and the less the tracking error (negative correlation). However, the direction of these relationships vary in an unpredictable fashion across recency measures, across CP task difficulty, and crew position. Moreover, the direction of the significant correlations is sometimes consistent with the above predictions and sometimes not. For example, months since last subcaliber fire is negatively correlated with error on the Hard tracking task (consistent) for tank commanders but it is positively correlated (inconsistent) for gunners.

Looking for the moment at the ASVAB composite scores, there is some evidence to indicate that CO and GT scores are statistically related to the computer tracking task, but only for gunners. For gunners, the higher the CO score, the more time on target on the average; this is significant only for the Hard tracking task. Additionally, GT scores for gunners are related in a similar fashion to time on target for both the moderate and hard Tasks.

In summary, the measures of computer tracking are related in a predictable fashion. There is evidence to suggest that the measures can be reduced to two for subsequent regression analyses--average time on target and average tracking error. However, for purposes of examining interrelationships among job samples for possible task differences, it is important to maintain the distinction between the Easy, Moderate, and Hard tasks.

More importantly, however, the fact that driver and loaders outperformed the tank commanders and gunners indicates there may be a negative transfer problem. Although tank commanders and gunners do not differ on the average in CT job sample performance, experience is negatively related to variation in the job sample measures within each of the tank commander and gunner samples.

4.3.3 Computer Target Engagement

Thirty-two tank commanders and 51 gunners contributed data to the Computer Target Engagement (CTE) job sample. Three tank commanders and two gunners did not participate because of scheduling problems.

Table 4-7 lists the 18 dependent measures which were derived from the Computer Target Engagement (CTE) job sample. There is a measure of accuracy (distance from the center of the reticle to the target) and time (elapsed) for each of four successive trial segments--the time during which the 3X target image was present (3X segment), the time from onset of the 10X target image to the initial press of the laser button (Laser 1 segment), the time from the initial press to the second press of the laser button (Laser 2 segment), and the time from either initial laser button press (multiple return bar absent) or second laser button press (multiple return bar present) to press of the fire button (Fire segment). Since the Laser 1, Laser 2, and Fire segments all represented time in which the 10X target image was present, the average error and average latency in the 10X segment was also computed. In addition, since it was noticed that the distribution of error scores across trials tended to be skewed, median as well as average errors were calculated. For a more complete description of the dependent measures, consult Appendix A.

TABLE 4-7. DEPENDENT MEASURES DERIVED FROM THE COMPUTER TARGET ENGAGEMENT JOB SAMPLE

Variable Name	Brief Description
PROC:ERROR	Laser Procedural Error; Failure to Lase when Multiple Return Bar Present or Lasing Twice when Multiple Return Bar Absent
3X:ERROR (AVG)	Average Distance Error (number pixels) at End of 3X Segment
3X:ERROR (MDN)	Median Distance Error (number pixels) at End of 3X Segment
L1:ERROR (AVG)	Average Distance Error (number pixels) at End of Laser 1 Segment
L1:ERROR (MDN)	Median Distance Error (number pixels) at End of Laser 1 Segment

TABLE 4-7. DEPENDENT MEASURES DERIVED FROM THE COMPUTER TARGET ENGAGEMENT JOB SAMPLE (continued)

Variable Name	Brief Description
L2:ERROR (AVG)	Average Distance Error (number pixels) at End of Laser 2 Segment
L2:ERROR (MDN)	Median Distance Error (number pixels) at End of Laser 2 Segment
F:ERROR (AVG)	Average Distance Error (number pixels) at End of Fire Segment
F:ERROR (MDN)	Median Distance Error (number pixels) at End of Fire Segment
10X:ERROR (AVG)	Average Distance Error (number pixels) of the L1, L2, and F Segments
10X:ERROR (MDN)	Average of the Median Distance Error (number pixels) of the L1, L2, and F Segments
3X:TIME	Average Time (sec) in 3X Segment
L1:TIME	Average Time (sec) in Laser 1 Segment
L2:TIME	Average Time (sec) in Laser 2 Segment
F:TIME	Average Time (sec) in Fire Segment
10X:TIME	Average Total Time (sec) in 10X Segment
TOT:TIME (AVG)	Average Time (sec) from Beginning to End of Trial
TOT:TIME (MDN)	Median Time (sec) from Beginning to End of Trial

Table C-8, in Appendix C, presents the intercorrelations among the CTE measures. As expected, there is a reasonably high positive correlation between the average and median distance error within each time segment, with the lowest correlation being +.558. This indicates that only one of the two measures should be used for subsequent (Phase II) analysis purposes. If the average and median distance measures have similar relationships with other variables, it would be best to select the average measures because of their mathematical properties.

Table C-8 also reveals high intercorrelations among distance error measures for the Laser 1, Laser 2, and Laser 3 time segments, with the correlations being higher for the average distance error measures. Since the 10X target image was present during each of these time segments, these high intercorrelations indicate that the size of the variable set can be reduced to that of 10X:ERROR. The high intercorrelations were not unexpected in that the target scene (10X slide) did not change, thereby not requiring the tank crewmen to change the position of the reticle over the target (except for minor adjustments).

One might expect that a tank crewmen's accuracy in laying the reticle on the 3X target to be highly correlated (positive) with his accuracy on the 10X target. That is not the case, however. Although there is some evidence to indicate that 3X:ERROR is significantly and positively correlated with measures of error during the 10X time segment (L1:ERROR, L2:ERROR, F:ERROR, 10X:ERROR), the correlations are low to moderate, with the highest correlation being +.351. Strangely the correlations between 3X:ERROR and measures of error during the 10X time segment are higher for tank commanders than gunners with the median error measures, whereas these correlations are higher for gunners than tank commanders with the average error measures. The fact that there is not a high correlation between 3X and 10X error is understandable when one examines the behavior of the tank crewmen on this task. Some tank crewmen adopted the strategy of initiating the 10X target scene once the 3X target was identified without being concerned for their accuracy with the 3X target. This would obviously tend to reduce the magnitude of the correlation.

As expected, the time measures are all positively correlated. However, the correlations among L1:TIME, L2:TIME, and F:TIME are all low and nonsignificant except in one case--the correlation between L1:TIME and F:TIME for gunners. When Laser 1, Laser 2, and Fire times are correlated with total time in 10X, the correlations are all significant, with those for Laser 1 time being much higher. The high correlation of L1:TIME with 10X:TIME indicates that most of the variation in time spent in 10X is the result of variation in the period from onset of the 10X visual scene to initial press of the laser button. Finally, 3X:TIME is positively and significantly correlated with both L1:TIME and 10X:TIME, with the correlation being much higher for gunners than tank commanders.

When examining the relationship between time and error measures, it is most meaningful to restrict the comparisons to those within a given time segment. The correlation between time and error are generally higher for the 3X and L1 time segments than they are for the L2 and Fire segments. The only significant correlations are those for the 3X and L1 segments. For the 3X and L1 time segments, the correlatons between time and error are all negative (with one exception). This indicates that the faster in which the tank crewmen responded, the more likely they were to make an error in laying the reticle on the target. For the Laser 2 and Fire segments, there is no consistent pattern among the correlations between time and error.

Table 4-8 presents the means and standard deviations for the tank commanders and gunners on each of the 18 dependent measures along with the results of

TABLE 4-8. COMPUTER TARGET ENGAGEMENT JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 18 DEPENDENT MEASURES

Dependent Measures		TC (n=32)	G (n=51)	F[1] Value	t[2] Value
PROC:ERROR	M	4.56	4.06		0.72
	SD	3.06	3.15	1.06	
3X:ERROR(AVG)	M	26.71	29.53		-1.08
	SD	10.25	12.37	1.45	
3X:ERROR(MDN)	M	16.77	15.98		0.31
	SD	10.96	11.29	1.06	
L1:ERROR(AVG)	M	12.00	14.04		-1.61
	SD	4.54	7.01	2.39*	
L1:ERROR(MDN)	M	6.13	6.76		-1.20
	SD	1.84	2.92	2.52**	
L2:ERROR(AVG)	M	14.50	16.70		-1.20
	SD	6.54	10.25	2.46**	
L2:ERROR(MDN)	M	9.28	10.65		-0.71
	SD	7.21	9.31	1.67	
F:ERROR(AVG)	M	12.45	14.59		-1.64
	SD	4.52	7.37	2.66**	
F:ERROR(MDN)	M	6.18	7.11		-1.87
	SD	1.67	2.90	3.00**	
10X:ERROR(AVG)	M	12.98	15.11		-1.52
	SD	4.88	7.85	2.59**	
10X:ERROR(MDN)	M	7.20	8.18		-1.14
	SD	3.22	4.58	2.02*	
3X:TIME	M	10.64	12.19		-1.44
	SD	4.15	5.11	1.51	
L1:TIME	M	12.83	12.83		0.00
	SD	4.41	4.63	1.10	
L2:TIME	M	1.69	1.62		0.22
	SD	1.27	1.37	1.17	

TABLE 4-8. COMPUTER TARGET ENGAGEMENT JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 18 DEPENDENT MEASURES (continued)

Dependent Measures		TC (n=32)	G (n=51)	F[1] Value	t[2] Value
F:TIME	M	2.66	2.71		-0.13
	SD	1.36	2.09	2.36*	
10X:TIME	M	16.13	1.6.17		-0.04
	SD	4.79	5.96	1.55	
TOT:TIME(AVG)	M	26.76	28.36		-0.75
	SD	7.52	10.37	1.90	
TOT:TIME(MDN)	M	24.36	25.89		-0.84
	SD	6.77	8.78	1.68	

* $p \leq .05$
** $p \leq .01$

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

the statistical tests comparing these two groups. The means and standard deviations for the drivers and loaders can be found in Appendix C (Table C-9). Several observations are noteworthy. First, the magnitude of the distance error appears quite larger. However, it must be remembered that the distance error represents the number of computer graphics pixels from the center of the reticle to the center of mass of the target. Given the size of the target imagery, an average of approximately four to six pixels probably represents a high degree of accuracy (a hit).

Secondly, the measures of average distance error are much larger than are the measures of median distance error. This confirms that the error measures are skewed across trials. Close inspection of the raw data reveals that the distance error is much greater on certain trials than others. There is a tendency for the distance error to be greater on early trials and on trials in which it appeared target detection is more difficult.

Thirdly, the accuracy from the center of the reticle to the target is much greater in 10X than it is in 3X. Two factors could be responsible for this finding--the smaller size targets in 3X and the strategy on the part of some tank crewmen to rapidly switch to 10X without concern for accuracy in 3X.

Finally, total time for target engagement (approximately 28 sec) is much longer than one would expect given the emphasis on speed of responding. In part, this may be due to the novelty of the procedures (lasing) and the emphasis placed upon accuracy as well as speed. As expected, most of the time spent in 10X involves the Laser 1 segment. Once the tank crewman locks on to the target and fires the laser, the remainder of the target engagement sequence is rapid. Interestingly, L1:TIME is slightly longer than 3X:TIME, suggesting the greater concern for accuracy in 10X than 3X. That is, as should be the case, subjects are taking more time in 10X to increase their accuracy.

The role experience should play in predicting differences between tank commanders and gunners is not clear cut. The tank commanders, simply based upon the longevity of their service, are probably more experienced in the behavior required in the CTE job sample. However, this may be counteracted by the recency of their practice. Because the gunners have had more recent practice in tasks which require behaviors involved in the CTE job sample, they may perform as well as, or even better, than the tank commanders. In addition, there is another element which would lead one to predict no difference between tank commanders and gunners--namely, the novelty of the task. The CTE job sample differs from hands-on target engagement in the M60A1 in two major respects--in the procedures required and in device fidelity. For example, in the M60A1 tank, there is no requirement to lase, and possibly release, to range to the target. In the tank, the reticle does not move off the center of the viewing area to acquire the target; rather, the reticle remains stationary in the center of the viewing area and it is the gun tube which moves. Lastly, the dynamics of the joystick controls differ from those of the tank's cadillac controls. As can be seen, the CTE job sample may be sufficiently different from target engagement in the tank itself so that differences in experience may not manifest themselves.

In spite of the fact that tank commanders and gunners may differ in the longevity of service and recency of practice, these differences are not evident in the performance measures of the CTE job sample. As can be seen in Table 4-8, there are no significant differences between tank commanders and gunners in their average level of performance on any of the 18 measures. Significant differences are manifested, however, in the variability of their performance on eight of the dependent variables, seven of which involve distance error. In each instance, there is greater variability among the gunners in their performance than among tank commanders. Since the variability of the measure affects the size of the correlation coefficient, care must be taken in interpreting differences between tank commanders and gunners in the significance of correlations with the CTE performance measures.

When drivers and loaders are considered, they should not perform as well as tank commanders and gunners because of their lack of practice at tasks requiring behaviors involved in the CTE job sample. That this is the case can be seen upon inspection of Table C-9 (Appendix C). Drivers and loaders have the largest average distance error in every case except two (median L1:ERROR and median L2:ERROR) and, in every case, take the longest time. These relatonships are significant, however, only for 3X:ERROR (AVG), 3X:TIME, and TOT:TIME (both AVG and MDN). Interestingly, drivers and loaders make fewer procedural errors (approximately one-half as many) than do tank commanders and gunners. It appears as if having had practice in established gunner procedures interferes with the performance of a task involving of a new procedural step of having to lase before firing.

Table C-10 (Appendix C) contains the correlations between the biographical measures and the CTE performance measures. Upon examining the interrelationships between measures of experience and measures of job sample performance, the following picture emerges. There is little evidence that the measures of experience are related to the measures of job sample performance for either task commanders or gunners. The correlations are low and, in general, nonsignificant.

If a prediction had been advanced, one would expect that distance errors would be negatively correlated with experience--i.e., the more experience, the less the error. There is a trend for the correlations to be in that direction for A1:TIME, A3:TIME, CP:TIME (gunners only), and ARMY:TIME (gunners only). For RANK, CP:TIME (tank commanders), and ARMY:TIME (tank commanders), the trend is in the opposite direction. However, in only six instances did the correlaton coefficient achieve significance.

With regard to the time measures, one would predict that more experienced tank crewmen would in all probability take less time (negative correlation). However, the direction of the correlations which are obtained between the measures of experience and the CTE time measures vary as a function of the measure of experience and crew position. For example, for A1:TIME and A3:TIME, the correlations are generally positive for gunners but negative for tank commanders. However, for RANK, CP:TIME, and ARMY:TIME, no clear picture emerges. In any case, the correlations are low and, in all but two instances, nonsignificant.

There are some interesting relationships which emerge with regard to measures of recency of training. For the VRT:MLAST measure, the longer since last vehicle recognition training, the greater the distance error. This relationship is significant, however, only for tank commanders. For gunners, there is statistical evidence that the longer since last CTT training, the greater the error in laying the reticle on the target. But, contrary to the finding for the gunners, the relationship between CTT:MLAST and the distance error measures for tank commanders is opposite in direction (and nonsignificant).

There is no statistical evidence that the ASVAB scores are, in general, related to measures of CTE performance. Only one correlation is significant. However, the pattern of the correlations of the ASVAB scores with CTE error measures is as expected. For both the CO and GT measures, the higher the ASVAB scores, the less the distance error from the center of the reticle to the target.

Finally, the correlations between the frequency with which computer games are played and the CTE measures are all nonsignificant with only one exception.

In summary, the pattern of correlations among the measures of performance for the CTE job sample is consistent with expectation. There is evidence that experience plays a role in CTE job sample performance in that drivers and loaders perform more poorly on this task than do tank commanders and gunners. However, experience is not statistically related to differences in CTE job sample performance, either between or within the tank commander and gunner subsamples.

Based upon the pattern of the results with the CTE job sample measures, the decision was made to focus attention on the following subset of five measures for subsequent presentation of the results:

- PROC:ERROR
- 3X:ERROR (AVG)
- 10X:ERROR (AVG)
- 3X:TIME
- 10X:TIME

Because of the high intercorrelation among the measures of 10X error, the decision was made to use 10X:ERROR rather than error during each individual 10X segment. Due to their mathematical properties, average error rather than median error is used. To create a situation parallel with that of error, 10X:TIME is employed rather than time during each 10X segment. Finally, since no other measure tapped procedural errors, it is included in the variable subset.

4.3.4 Tank Commander Decision Making

Two dependent measures were obtained on the Tank Commander Decision Making (TCD) job sample--the number of correct decisions (D:CORR) and the time to reach a decision (D:TIME). Appendix A contains the specifics regarding the measurement of these parameters.

Table C-11 in Appendix C presents the correlation between the two TCD measures for 35 tank commanders and 53 gunners who participated in the TCD job sample. For both tank commanders and gunners, the correlation between the time to reach a decision and number correct is essentially zero. This lack of a relationship between time and correctness is not unexpected and can be explained with aid of Table 4-9. Note in Table 4-9 that the average decision time is short and has a very small standard deviation for both tank commanders and gunners. This suggests that the tank crewmen emphasized time at the expense of accuracy. Responding quickly will have no effect on the correctness of those who are well practiced, but it will increase the likelihood of errors for those who do not readily know the answer. Therefore, very small differences in response times are associated with large differences in number correct.

TABLE 4-9. TC DECISION MAKING JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 2 DEPENDENT MEASURES

Dependent Measures		TC (n=35)	G (n=53)	F^1 Value	t^2 Value
D:CORR	M	15.54	15.66		-0.18
	SD	3.15	2.95	1.14	
D:TIME	M	3.77	3.95		-0.88
	SD	0.82	1.01	1.52	

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

The means and standard deviations for the tank commanders and gunners on the two TCD measures are shown in Table 4-9 along with results of the statistical comparison of these two groups. It was expected that since the decision as to which target to engagement is the tank commander's responsibility, they should do better on this task than the gunners. Certainly tank commanders should have more experience and be much more practiced at this task than is the case for gunners. Therefore, it was predicted that tank commanders would have more correct and a shorter decision time than gunners. However, there is no evidence that this is the case in that the tank commanders and gunners have almost an identical number correct and differ very slightly in decision time. Obviously, these differences are nonsignificant.

There are at least two alternative explanations for the failure to find differences between tank commanders and gunners on this task. It may be that although target engagement decisions are measured in this job sample, some general decision making aptitude is involved. Since both tank commanders and gunners are involved in making certain cognitive decisions, one would not necessarily expect a difference if the TCD job sample is a test of general decision making ability. Alternatively, maybe tank commanders have evolved their own strategies for engagement which run counter to the rules of engagement. In a multiple engagement situation, it may be more important to some tank commanders to open fire on any reasonable target rather than to waste time deciding which target to engage. This would be particularly true when differences in the targets are slight (as was the case on several trials).

That the latter is the case is suggested by the correlations of the measures of experience with decision making correctness. As indicated in Table C-12 (Appendix C), the correlation of four of the five experience measures (RANK, ARMY-TIME, A1:TIME, CP:TIME) with D:CORR for tank commanders is negative; that is, the greater the length of service (i.e., the more experience), the fewer the number correct. This relationship is significant, however, only for ARMY:TIME. Although the sign of the correlation is also negative for two of the five relationships for gunners, those correlations are essentially zero.

Further inspection of Table C-12 reveals that none of the correlations with D:TIME are significant. This is not unexpected in that there is very little variability in the time measure. Additionally, the correlation of the job sample measures with measures of recency of training, ASVAB scores, and the frequency with which crewmen play computer games are all low and nonsignificant with only one exception--that between CTT:MLAST and D:CORR. However, the direction of the significant correlation does not make any sense.

In summary, the lack of a significant difference between tank commanders and gunners on the TCD measures is not what one would have expected based upon the presumed difference in experience of the two groups. Although the lack of a significant difference was not predicted, it does make sense when certain post hoc reasoning is applied.

4.3.5 Hands-On Gun Laying

Eighty-six of the 88 tank crewmen participated in the Hands-On Gun Laying (HGL) job sample. One tank commander and one gunner did not take part due to scheduling problems.

Two dependent measures were gathered for the HGL job sample. These were gun laying error (GL:ERROR) and time to lay the gun (GL:TIME). For a more complete description of these two measures, see Appendix A.

Table C-13 (Appendix C) contains the correlation between time to lay the gun and gun laying error. As seen in Table C-13, gun laying time and error are not significantly related, with the correlation being negative for tank commanders and positive (almost zero) for gunners. These results are perfectly understandable when one considers the following. The emphasis in gun laying is on speed and not accuracy. The strategy is to place the gun tube, as rapidly as possible, in the general direction of the target so that the gunner can take control. As long as the target is somewhere in the gunner's sight, the gunner can take over. Therefore, gun laying time is sometimes associated with a small error and sometimes with a large error, thus resulting in a nonsignificant correlation. For those individuals skilled in manipulating the TC power handle and the gun tube (i.e., the tank commander), taking a slightly larger time should increase accuracy. Thus, a positive correlation would result. However, for those unskilled in gun tube manipulation (i.e., the gunners), taking more time will not necessarily help them (i.e., zero correlation).

As part of their assigned responsibilities, tank commanders routinely practice laying the gun. On the other hand, given it is outside their domain of responsibility, gunners do not receive much training in gun laying nor do they regularly practice it. Therefore, based upon this difference in experience and training, tank commanders should clearly perform at a higher level than gunners. Since the emphasis in gun laying is on speed rather than accuracy, tank commanders should take less time to lay the gun than gunners but not necessarily be any more accurate. Results of the statistical analyses (cf, Table 4-10) confirm this prediction. As indicated in Table 4-10, tank commanders take significantly less time than gunners to lay the gun but do not differ from the gunners in gun laying error.

TABLE 4-10. HANDS-ON GUN LAYING JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON TWO DEPENDENT MEASURES

Dependent Measures		TC (n=34)	G (n=50)	F^1 Value	t^2 Value
GL:TIME	M	6.61	8.26		-3.37***
	SD	1.83	2.43	1.76	
GL:ERROR	M	6.88	7.23		-0.62
	SD	2.53	2.44	1.08	

***$p \leq .001$

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

Table C-14 (Appendix C) presents the correlations of the HGL performance measures with the measures of experience, recency of training, and ASVAB test scores. Consistent with the above results, time to lay the gun is negatively, and nonsignificantly, correlated with four of the five experience measures (RANK, ARMY:TIME, A1:TIME, CP:TIME) for tank commanders. For gunners, the correlations of gun laying time with the experience measures are very low, nonsignificant, and not in a consistent direction across measures. Furthermore, there is no consistent pattern in the relationship between the experience measures and gun laying error, the correlations being in general low and nonsignificant. The only correlation which is significant, that between GL:ERROR and ARMY:TIME, is in the direction (positive) opposite to expectation.

The correlations of the HGL measures with the ASVAB composite scores fit a predictable pattern. In general, the higher the CO and GT scores, the less the time and error in laying the gun. These relationships, however, are significant in only one case--namely, the correlation between CO scores and gun laying error for gunners.

The correlations relating measures of recency of training and frequency of playing computer games with the HGL measures are all nonsignificant.

In summary, experience is definitely a factor in the hands-on gun laying job sample in that tank commanders and gunners differed in performance on that task. In spite of this difference, there is no statistical evidence that experience is related to the variation of the HGL measures within the tank commander and gunner subsamples.

4.3.6 Hands-On Tracking

Table 4-11 lists the six dependent measures which were obtained for the 34 tank commanders and 52 gunners who participated in the Hands-On Tracking (HT) job sample. There are two measures for each of the two crew positions--number of hits and the distance tracked. In addition, the number of hits and distance tracked were averaged across the two positions to yield two measures--TOT:HITS and TOT:DIST. See Appendix A for a more complete description of these measures.

TABLE 4-11. DEPENDENT MEASURES DERIVED FROM THE HANDS-ON TRACKING JOB SAMPLE

Variable Name	Brief Description
TOT:HITS	Number hits averaged across the TC and gunner station.
TOT:DIST	Distance (inches) tracked averaged across the TC and gunner station.
TC:HITS	Number hits at the TC station.
TC:DIST	Distance (inches) tracked at the TC station.
G:HITS	Number hits at the gunner station.
G:DIST	Distance (inches) tracked at the gunner station.

Inspection of Table C-15 (Appendix C) reveals that all dependent measures were significantly intercorrelated. As expected, distance tracked and number of hits are negatively correlated at each crewstation for both tank commanders and gunners--namely, as the distance tracked increases, the number of hits decreases. This indicates that as the tracking speed increases, accuracy decreases.

The number of hits at the two crewstations and the distance traveled at the two stations are positively correlated, with the correlations being higher for the distance measure than the hits measure (cf, Table C-15). If one assumes some underlying tracking skill, these positive correlations are not unexpected. The lower correlation for the hits measure is understandable in that tracking is not normally accomplished at the TC station. Neither tank commanders nor gunners would have developed the proficiency to track accurately at the TC station. Therefore, in spite of some underlying tracking skill, accuracy at the two stations would not necessarily be highly correlated. However, given the instructions to track as accurately and as fast as possible, crewmen may have made the attempt to track at approximately the same rate at both crewstations, thus resulting in a high correlation.

The fact that the number of hits at the two crewstations is not highly correlated suggests that they not be averaged. Furthermore, in spite of their high intercorrelation, the two distance measures should not be averaged in subsequent Phase II analyses in order to create a situation parallel to that for the two hit measures.

Table 4-12 contains the means and standard deviations for the tank commanders and gunners for the six HT measures, as well as the results of the statistical analyses comparing the two groups. As can be seen, for both tank commanders and gunners, tracking proficiency is greater at the gunner's station--namely, more hits are made and a greater distance is tracked at the gunners as opposed to the TC station. More importantly, however, Table 4-12 shows that the tank commanders and gunners do not statistically differ on any of the six measures.

In advance of collecting the data, it was reasonable to assume that the tank commander should track as equally well or better than the gunners at the TC station. In that tank commanders have had experience at the TC station, whereas gunners have not, one might predict that tank commanders would outperform the gunners at the TC station. On the other hand, since tracking is a relatively novel task at the TC station, there was no reason to expect that experience would be an advantage. Obviously, the fact that there was a lack of significance on either measure at the TC station is in accordance with the second expectation.

At the gunner's station, gunners should have tracked as well or better than the tank commanders. In that tank commanders and gunners have had experience tracking at the gunner's station, they should not differ. However, given the recency of the gunner's practice, one might have expected them to perform better than the tank commanders at the gunner station. The fact that there were no significant differences between tank commanders and gunners at the gunner's station supports the first contention.

TABLE 4-12. HANDS-ON TRACKING JOB SAMPLE: MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 6 DEPENDENT MEASURES

Dependent Measures		TC (n=34)	G (n=52)	F[1] Value	t[2] Value
TOT:HITS	M	43.53	42.12		1.12
	SD	5.57	5.84	1.10	
TOT:DIST	M	230.04	198.48		1.90
	SD	74.40	75.98	1.04	
TC:HITS	M	36.21	33.09		1.88
	SD	7.93	7.29	1.18	
TC:DIST	M	220.27	190.24		1.80
	SD	72.76	77.19	1.13	
G:HITS	M	50.85	51.16		-0.27
	SD	4.95	5.12	1.07	
G:DIST	M	239.80	206.71		1.87
	SD	81.49	79.36	1.05	

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

If experience is a factor in HT job sample performance, one would expect the measures of experience to be positively related to the number of hits at the gunner's station and negatively related at the TC station. Prediction of the direction of the relationship between experience and tracking distance is not clear cut. It may be that the more experience, the more likely one is to slow down to maximize the number of hits and, thus, a negative correlation would result. On the other hand, if experience lends to increased accuracy, the more experienced crewman may be able to track further; thus, a positive correlation would be expected.

Table C-16 presents the intercorrelations between the measures of experience and HT job sample performance. None of the correlations between number of hits and measures of experience are significant. Consistent with expectation, four of the five experience measures for tank commanders and four of the five experience measures for gunners (not necessarily the same ones) are negatively related to the number of hits at the TC station. Somewhat disappointingly, the correlations between the number of hits and the measures of experience at the gunner's station are not all positive and follow no consistent pattern.

The only significant correlations which were obtained with the measures of experience were those with tracking distance. In all experience measures but one (A3:TIME), the correlations were positive for both tank commanders and gunners indicating that the more the experience, the greater the distance tracked.

None of the correlations with the recency of experience measures, the ASVAB composite scores, or the frequency of playing computer games are significant. Moreover, the directions of the correlations vary across measures, crew-station, and subsample and, therefore, fit no predictable pattern.

In summary, as with the other job samples, interrelationships among the HT measures fit a predictable pattern. Moreover, if experience is a factor, it does not manifest itself in HT performance either in differences between tank commanders and gunners or in relationships between measures of experience and number of hits. The only evidence for experience as a factor in HT performance comes from the relationships between the experience measures and the distance tracked. Here, the more the experience, the greater the distance tracked.

4.3.7 Hands-On Target Engagement

Eighty-three tank crewmen took part in the Hands-On Target Engagement (HTE) job sample. Due to scheduling problems, one tank commander and four gunners did not particpate in this task.

Four dependent measures were derived from the HTE job sample. They are listed in Table 4-13. For a more complete description of the measures, consult Appendix C. Although the reason for measuring TOT:HITS and G:TIME is self-evident, the reason for measuring TC:TIME is worthy of note. TC:TIME represents the time from onset of a trial to the point at which the subject occupying the gunner's position takes control. This time period includes variable time which is not under the subject's direct control--namely, the time it takes the person occupying the TC station (a confederate of experimenter) to detect the target and lay the gun. Although a well practiced tank commander was always used as the confederate, his gun laying time would not be constant across subjects. Moreover, different confederates were used for different subjects. Despite this noise in the TC:TIME variable, it does contain a systematic component related to the subject's performance at the gunner station. The systematic component in the TC:TIME measure represents the time it takes the gunner to detect the target in his field of view and take control of the turret.

Table C-17 in Appendix C contains the intercorrelation among the HTE measures. As necessarily must be the case, the two time measures are significantly correlated with the total time. Interestingly, the correlations with total engagement time are much higher (approximately +.930) for TC:TIME than for G:TIME. This indicated that variation in TC:TIME is much

TABLE 4-13. DEPENDENT MEASURES DERIVED FROM THE HANDS-ON TRACKING JOB SAMPLE

Variable Name	Brief Description
TOT:HITS	Total number of hits in 15 trials
TOT:TIME	Average total time (sec) from onset of a trial to press of the gunner's trigger
TC:TIME	Average time (sec) from onset of a trial to point at which TC removes hands from TC power handle
G:TIME	Average time (sec) from point at which TC removes hands from TC power handle to press of gunner's trigger

more responsible for variations in total engagement time than is G:TIME. Additionally, the high correlations of the component times with TOT:TIME points to the fact that they cannot be utilized simultaneously in subsequent Phase II analyses. Given the low and nonsignificant correlations between TC:TIME and G:TIME, it probably would be best to use the component time measures rather than total engagement times in that the component measures each convey different information.

The correlations between number of hits and the two component time measures (TC:TIME, G:TIME) are all nonsignificant. For TC:TIME, one would not necessarily expect a relationship. However, for G:TIME, it might be possible to argue that if a gunner took slightly more time, he could increase his accuracy and, thus, a positive correction with number of hits would result. But when it is remembered that the emphasis in tank gunnery is also on speed of responding, one could argue that responding quickly would increase the likelihood of an error. As a consequence, a negative relationship with number correct would occur. The fact that nonsignificant correlations were obtained probably represents these two counteracting tendencies at work.

Table 4-14 presents the means and standard deviations for tank commanders and gunners on each of four HTE measures, as well as the results of the statistical comparisons of the two groups. Based upon the role of experience in these two groups, two different outcomes are plausible. On the one hand, since both tank commanders have had experience and practice at the gunner's station, they should perform equally well. On the other hand, given the recency of the gunners' practice, they may outperform the tank

TABLE 4-14. HANDS-ON TARGET ENGAGEMENT JOB SAMPLE: MEAN, STANDARD DEVIATIONS, AND SIGNIFICANCE FOR TANK COMMANDERS (TC) AND GUNNERS (G) ON 4 DEPENDENT MEASURES

Dependent Measures		TC (n=34)	G (n=49)	F^1 Value	t^2 Value
TOT:HITS	M	11.44	10.84		1.25
	SD	2.31	2.06	1.27	
TOT:TIME	M	13.44	14.21		-0.88
	SD	3.74	3.99	1.14	
TC:TIME	M	8.77	9.54		-1.04
	SD	2.97	3.56	1.44	
G:TIME	M	4.67	4.67		0.00
	SD	1.49	1.46	1.03	

[1]Tests significance of homogeneity of variance assumption.

[2]Tests significance of difference between means. Unpooled variance estimate used when significant heterogeneity of variance.

commanders. The results support the first prediction in that there were no significant differences between the two groups on any of the measures.

If experience is a factor in the HTE job sample, then within each subsample measures of experience should be positively correlated with the number of hits. This is the case for four of the five measures of experience (Rank, ARMY:TIME, A1:TIME, CP:TIME). However, as shown in Table C-18 (Appendix C), none of these correlations are statistically significant.

Prediction of the direction of the relationship between experience and gunner time is not clear cut. Perhaps more experienced tank crewmen have learned that by taking slightly more time, they can increase their accuracy. Alternatively, the more the experience, the greater the ability to fire quickly without a loss of accuracy. The fact that correlations of G:TIME with the measures of experience are all nonsignificant and do not follow a consistent pattern suggests that both tendencies are in operation.

The correlations of the HTE measures with the measures of recency of practice, the ASVAB composite scores, and computer game frequency are also nonsignificant and do not follow a consistent interpretable pattern.

In summary, the interrelationship among the HTE measures is also plausible. Experience, if it is a factor in HTE performance, does not manifest itself in group difference between tank commanders and gunners or in correlations with measures of experience.

4.3.8 Summary of Within-Job Sample Relationships

The summary of the within-job sample relationships is organized around the three purposes advanced at the outset of Section 4.3.

4.3.8.1 Degree of Correlation and Redundancy Among Measures Within the Job Samples

As has been seen, there is a fairly high degree of intercorrelation among some of the measures within given job samples indicating redundancy among the measures. Based upon the redundancy among the measures, the pattern of intercorrelation, and a logical analysis of the reasons for inclusion of the measures, it was decided to reduce the variable list for subsequent presentation of the results to the subset of measures listed in Table 4-15. This reduced subset of variables is hereafter identified as the primary measures.

For purposes of regression analysis in Phase II of the present investigation, it is possible to reduce the list of variables even further. Because of the high intercorrelation among the measures for the computer tracking task, it is possible to substitute average time on target (AVG:TOT) and average tracking error (AVG:ERROR) for the comparable measures on the separate CT tasks.

TABLE 4-15. REDUCED SUBSET OF DEPENDENT MEASURES IDENTIFIED AS THE PRIMARY MEASURES

Job Sample	Primary Measures
Computer Panel (CP)	ECD:CORR ECD:TIME CST:CORR CST:TIME
Computer Tracking (CT)	EASY:TOT EASY:ERROR MOD:TOT MOD:ERROR HARD:TOT HARD:ERROR
Computer Target Engagement (CTE)	PROC:ERROR 3X:ERROR(AVG) 3X:TIME 10X:ERROR(AVG) 10X:TIME

TABLE 4-15. REDUCED SUBSET OF DEPENDENT MEASURES IDENTIFIED AS THE PRIMARY MEASURES (continued)

Job Sample	Primary Measures
TC Decision Making (TCD)	D:CORR D:TIME
Hands-On Gun Laying (HGL)	GL:TIME GL:ERROR
Hands-On Tracking (HT)	TC:HITS TC:DIST G:HITS G:DIST
Hands-On Target Engagement (HTE)	TOT:HITS TC:TIME G:TIME

4.3.8.2 Construct Validity of the Job Samples

In assessing the construct validity of the job samples, logical deductions were made about the direction of relationships between job sample measures and then compared to the obtained correlations to determine if the presumed relationships occurred. There was good evidence for the construct validity of individual job samples in that the relationships which emerged among job sample measures in all cases either fit the predicted pattern or were perfectly understandable given knowledge of tank gunnery and human performance measurement.

4.3.8.3 Role of Experience in Job Sample Performance

Table 4-16 presents a summary of the evidence regarding the role of experience as a factor in job sample performance. In preparing this table, it was assumed that the more experienced tank crewman should achieve a higher degree of correctness and accuracy on the job samples as well as respond in a shorter period of time. This presumed relationship constituted the "expected direction" for purposes of table construction.

As Table 4-16 indicates, presumed differences in the experience level of tank commanders and gunners resulted in significant differences in job sample performance in only one case--Hands-On Gun Laying. Furthermore, when differences in experience within tank crewmen subsamples are considered, there is statistical evidence for experience being a factor only for the Computer Tracking job sample. In that job sample, quite opposite to expectation, experience is in general negatively related to performance for both tank commanders and gunners; that is, higher experience levels are

TABLE 4-16. SUMMARY OF EVIDENCE FOR EXPERIENCE AS FACTOR IN JOB SAMPLE PERFORMANCE

	Experience Between TC and Gunner Subsamples	Experience Within Tank Commander Subsample*			Experience Within Gunner Subsample*		
Job Sample	Significant Differences	Percentage of Bivariate Correlations in Expected Direction	Percentage of Bivariate Correlations Which are Significant	Above Chance Level	Percentage of Bivariate Correlations in Expected Direction	Percentage of Bivariate Correlations Which are Significant	Above Chance Level
Computer Panel	No	35(20)	5	No	55(20)	0	No
Computer Tracking	No	3(30)	20	Yes	23(30)	6	No
Computer Target Engagement	No	52(25)	8	No	44(25)	4	No
TC Decision Making	No	20(10)	10	No	80(10)	0	No
Hands-On Gun Laying	Yes	60(10)	10	No	40(10)	0	No
Hands-On Tracking	No	50(20)	0	No	50(20)	5	No
Hands-On Target Engagement	No	60(15)	13	No	60(15)	0	No

*Experience here refers to correlations with the biographical experience measures (Rank, Army:Time, A1:Time, A3:Time, CP:Time).

associated with less time on target and greater tracking error. The negative relationship is above chance level only for tank commanders, however.

Although not statistically above chance level, there is consistency in the direction of the relationship between the measures of experience and the TC Decision Making job sample performance measures. The relationships with the experience measures are opposite in direction, however, for tank commanders and gunners. For gunners, the expected positive relationship emerges. But for tank commanders, there is a trend for the more experienced crewmen to make fewer correct decisions and take more time to reach a decision.

In the other job samples, measures of experience are not consistently related to performance differences within the tank commander and gunner subsamples.

The fact that there is a lack of statistical evidence for experience as a factor in job sample performance at least makes tenable the assumption that one is measuring aptitudes. However, as discussed in detail in Section 4.1, one cannot definitely affirm that differences in aptitudes are responsible for differences in job sample performance.

4.4 ANALYSIS OF THE INTERRELATIONSHIPS AMONG JOB SAMPLES

The interrelationship among measures of different job samples was investigated to establish the degree of intercorrelation among job sample measures and to determine the plausibility of the relationships which emerged.

The analysis of the interrelationships among the job samples proceeded at two levels--at the level of bivariate (zero order) correlational analysis and at the level of canonical correlational analysis. The present section examines the results of each of these analyses.

At the level of bivariate correlational analyses, the interrelationships among the job samples were examined in a pairwise fashion--namely, two job samples at a time. Given seven job samples, there are twenty-one relationships to be examined. To facilitate exposition of the results, the relationships were organized into three groupings. These include the interrelationships involving just the computer job samples, those involving just the hands-on job samples, and those involving a comparison of the computer with the hands-on job samples. The Tank Commander Decision Making (TCD) job sample, although it could have been performed anywhere, was conducted on the tank and is, therefore, considered part of the hands-on job samples.

The bivariate correlations which resulted from the pairwise comparisons are tabled in Appendix D, Tables D-1 to D-21. In evaluating the bivariate correlations, attention is focused on the primary (reduced) set of measures for each job sample (see Section 4.3). To facilitate examination of the tables in Appendix D, the relationships among the primary variables are boxed in. When warranted, significant correlations for the other measures are considered as well.

Furthermore, in examining the interrelationships as to their plausibility, comparisons are restricted to those involving similar measures, namely, accuracy (correctness) with accuracy and speed (time) with speed. There is no a priori reason to examine the relationships between measures of accuracy on one task and the measures of speed (time) on another. In absence of information to the contrary, it is most plausible to assume that measures of accuracy (correctness) would be positively related as would measures of speed (time).

In anticipation of the results which are considered in detail below, the results, in general, show a low and nonsignificant level of relationship between the measures of different job samples. This supports the contention that the job samples tap different aspects of the target engagement sequence and also indicates an ideal measurement situation for later regression analysis.

4.4.1 Computer Job Samples: Bivariate Relationships

4.4.1.1 Computer Panel vs Computer Tracking

In making comparisons between the Computer Panel (CP) sample and the other job samples, there is no a priori reason to expect any significant relationships. The CP job sample is the only one which does not have a heavy emphasis on the speed of responding. Furthermore, contrary to what is true for the other job samples, the CP job sample does not have a heavy psychomotor performance component. Rather, the CP job sample taps more of a cognitive ability--the ability to assimilate and correctly utilize new procedures.

As expected, there is little evidence to indicate that the measures of performance on the CP job sample are related to the measures of performance on the Computer Tracking (CT) job sample. As Table D-1 (Appendix D) indicates, significance is obtained in only two of the 24 cells involving comparison of the major variables (average measures excluded). The number of significant relationships do not exceed chance expection ($p = .223$). Of the two significant relationships, only one is plausible--that between ECD:CORR and HARD:ERROR. As the numbers correct on the Enter/Check Data CP task increases, the magnitude of the tracking error in the Hard CT task decreases.

In spite of the lack of significance, the direction of the obtained relationships were examined as to their plausibility. Since the CT task has no time or speed measure, attention was focused upon comparisons involving number correct on the CP tasks with time-on-target and tracking error for the CT task. For tank commanders, a generally consistent and interpretable pattern emerges. Namely, with only one exception, there is a trend for a greater number correct on the CP tasks to be associated with a longer time-on-target and a smaller tracking error. There is no clear trend, however, for gunners.

4.4.1.2 Computer Panel vs Computer Target Engagement

Based upon the nature of the Computer Panel (CP) job sample, only one relationship with the Computer Target Engagement (CTE) job sample would have been predicted in advance of collecting the data. Since the CP job sample is basically a procedural task, there should be a relationship between the number correct on the CP tasks and the number of procedural errors on the CTE job sample. More specifically, as the number correct on the CP tasks increases, the number of procedural errors on the CTE job sample should decrease. Inspection of Table D-2 (Appendix D) reveals that the expected negative relationship occurs for both CP measures (ECD:CORR, CST:CORR) and for both tank commanders and gunners. However, the relationship is significant in only one case--that for ECD:CORR with the combined subsamples.

Examination of Table D-2 shows that very few of the correlations relating the CP and CTE job samples are significant. When attention is focused upon the major variables for the two job samples, there is no evidence that the number of significant correlations exceed chance expectation. Out of twenty major relationships, only two correlations were significant for gunners, two for tank commanders, and two for the combined subsamples.

The lack of significant relationships between the CP and CTE job sample is not unexpected in that the measures on the two tasks are completely different. The number correct on an untimed task involving application of new procedures (CP job sample) bears no resemblance to a measure of psychomotor skill (distance error) on an untimed task (CTE job sample). Furthermore, time to complete a task (CP job sample) probably taps different behavior than does response latency (CTE job sample).

Despite the lack of significance, the pattern of interrelationships was examined as to its plausibility and consistency. If plausible relationships exist, one should find a negative correlation between the number correct on the CP tasks (ECD:CORR, CST:CORR) and the distance error measures [3X:ERROR(AVG), 10X:ERROR(AVG)] on the CTE task. For gunners, the expected relationship emerges in three out of the four cases, whereas for tank commanders, there is no consistent relationship.

A prediction with regard to the direction of relationship between the measures of time on the two tasks is difficult to advance. A positive and a negative correlation are equally plausible. On the one hand, the better tank crewmen (as indexed by short CTE times) may take longer on an untimed task to achieve higher accuracy. On the other hand, the better tank crewmen may respond rapidly on all tasks in that the emphasis in tank gunnery is on speed. In any case, consistent but opposite relationships emerge for the tank commanders and gunners. For tank commanders, there is a negative relationship between CP completion time and the length of the CTE time segments (with one exception), whereas for gunners, the relationship is positive in all cases.

4.4.1.3 Computer Tracking vs Computer Target Engagement

Although the Computer Tracking (CT) and Computer Target Engagement (CTE) job samples both require psychomotor skill, they have different performance components. First, the CTE job sample involves moving the reticle to the center of mass of a stationary target, whereas the CT task involves attempting to maintain the center of the reticle on the center of mass of a moving target. Secondly, the CTE task has a heavy cognitive component which is absent in the CT job sample. In the CTE task, the tank crewmen must learn and apply new procedures (i.e., lasing) and make decisions as to whether or not to lase.

Given these important task differences, there is no reason to necessarily expect a statistical relationship between the CT and CTE job samples. However, if plausible relationships exist, the CTE distance error measures (3X:ERROR, 10X:ERROR) should be negatively related to time-on-target and positively related to tracking error.

Table D-3 (Appendix D) contains the bivariate (zero order) correlations relating the measures of the CT and CTE job samples. The results indicate that six of the 30 possible correlations among the major variables are significant for gunners, whereas only one is significant for tank commanders. The number of significant correlations for gunners exceeds expectation based upon chance ($p = .003$).

Inspection of Table D-3 reveals that all the significant correlations for gunners deal with a cross comparison [i.e., the relationship between measures of accuracy on the CT task and a measure of time (3X:TIME) on the CTE job sample.] There is no a priori reason to expect such a relationship. Nevertheless, in all tracking tasks (Easy, Moderate, and Hard), the greater the time-on-target, the less the time spent in the 3X segment in the CTE task. Conversely, the greater the tracking error, the greater the time spent in 3X. There is a trend for the results to be in the same direction for tank commanders and for 10X:TIME, but obviously they are not as strong. These results, although not predictable, are plausible if one assumes some general underlying gunnery aptitude. If such an aptitude exists, then doing better in one task (more time-on-target and less tracking error) would be related to doing better at another task (faster response time).

Correlations between measures of accuracy on the two tasks are generally low and, with only one exception, are nonsignificant. Inspection of Table D-3 reveals that the expected pattern of relationships exist for 10X:ERROR(AVG), but not for 3X:ERROR(AVG). As can be seen, as the distance error in 10X decreases, there is a tendency for tracking error to decrease and time-on-target to increase. This relationship holds for both tank commanders and gunners. The lack of a consistent pattern between the measures of accuracy on the two tasks during the 3X time segment is understandable when it is remembered that during the 3X time segments, some tank crewmen are not concerned with their accuracy (cf, Section 4.3.3).

4.4.2 Hands-On Job Samples: Bivariate Relationships

4.4.2.1 TC Decision Making vs Hands-On Gun Laying

The Tank Commander Decision Making (TCD) job sample is primarily a cognitive task, whereas the Hands-On Gun Laying (HGL) job sample is basically a psychomotor task. On this basis, there is no reason to expect that the measures on the two tasks would necessarily be correlated, particularly in the case of accuracy.

That there would be no relationship between the measures of accuracy on the two tasks seems particularly plausible when it is remembered that the concern in gun laying is on speed and not precision of the gun lay. Rather, the concern is in getting the gun tube pointed in the general direction of the target as rapidly as possible. Although it can be measured, the accuracy of the gun lay would bear little relationship to gunnery skills and, therefore, not necessarily relate to decision accuracy. However, given the emphasis on speed of responding in both tasks, it seems plausible that if a relationship exists, time to make a decision should be positively related to gun laying time. Moreover, this relationship should be stronger for tank commanders than gunners since both job samples represent tasks rarely practiced by gunners.

Table D-4 indicates that the above analysis is supported by the data. More specifically, D:TIME and GL:TIME are positively correlated for both tank commanders and gunners with the relationship being significant only for tank commanders. On the other hand, the two measures of accuracy are not significantly correlated and do not exhibit the same direction of relationship for tank commanders and gunners. The direction of the correlation for gunners, however, is in accordance with what would have been expected had a prediction been advanced. Namely, the number correct decisions is negatively related to gun laying error.

4.4.2.2 TC Decision Making vs Hands-On Tracking

The same logic which was used in the last section can be applied to the relationship between the measures of performance for the TC Decision Making (TCD) and Hands-On Tracking (HT) job samples. On the basis that the two job samples represent different skills (one cognitive and the other psychomotor), there is no reason to predict that the measures of accuracy on the two tasks (number of correct decisions and number of hits in the tracking task) or the measures of speed on the two tasks (decision time and distance tracked) would necessarily be correlated. Furthermore, since tracking is not normally accomplished from the TC station, it is less likely to find relationships (if they exist) at the TC station.

Table D-5 (Appendix D) presents the bivariate correlations between the measures of performance on the TCD and HT job samples. There are six significant correlations, two each for the task commanders, gunners, and combined subsamples. The number of significant correlations approaches significance in each case ($p = .051$). All the significant correlations involve the tracking distance measures.

The correlations between tracking distance and decision correctness and between tracking distance and decision time follow a consistent pattern (with only one exception). As might be expected, decision time and the tracking distance are inversely related for both tank commanders and gunners. This relationship between decision correctness and tracking speed is significant at both the TC and gunner station but only for gunners.

Although not predictable in advance of collecting the data, decision correctness and tracking distance are inversely related. Again, these correlations with tracking distance are significant at both crewstations, but this time only for tank commanders. This relationship makes sense in that the better tank crewmen (as indexed by the number correct decisions and time to reach a decision) probably slow down their tracking speed to increase the number of hits on the tracking task. This would even be more likely at the TC station where the novelty of the task would really require the tank crewmen to slow down to achieve any degree of accuracy. The discrepancy in the differential significance for tank commanders and gunners is understandable if one assumes that decision correctness is a better index of tank commander skill and speed of responding (decision time) is a better index of gunner skill.

As expected, none of the correlations relating the measure of accuracy on the two tasks are significant. However, in all cases, the direction of the relationships are in accordance with what would have been expected had predictions been advanced. Namely, decision correctness and the number of tracking hits are positively related for both tank commanders and gunners.

4.4.2.3 TC Decision Making vs Hands-On Target Engagement

It is not unreasonable to expect a relationship between the Tank Commander Decision Making (TCD) and the Hands-On Target Engagement (HTE) job samples in that target engagement involves an element of decision making. However, the TCD task is primarily cognitive and the HTE is primarily psychomotor. Moreover, the types of decisions required in the TCD task are not involved in the HTE tasks. Therefore, if any correlations exist, they are likely to be low. If there are any significant relationships, they are more likely to occur with the time measures in that the measures bear a closer resemblance than do the accuracy measures. Accuracy in moving the reticle to the center of the target probably does not require the same component skills as does accuracy in making a decision.

Table D-6 (Appendix D) reveals that there is some statistical evidence for a relationship between the TCD and HTE job samples but only for the combined subsamples. None of the correlations relating the major variables achieve significance for either tank commanders or gunners. As expected, the correlations are in general low. However, the two significant correlations for the combined subsamples exceed the number expected based upon chance ($p = .031$).

As expected, the correlations relating decision time to the two HTE time measures are all positive, with the correlations being much higher for G:TIME. Despite the same relative magnitude of the correlations for tank commanders and gunners, the relationship between D:TIME and G:TIME is significant only for the combined subsamples. The lower correlations (almost zero) between D:TIME and TC:TIME were not unexpected in that TC:TIME has a large component of variability not under the subject's control (i.e., the time which it takes the confederate to lay the gun).

Although nonsignificant, the direction of the correlations between decision correctness and the number of hits is opposite to expectation. Those tank crewmen who make more correct decisions have fewer hits during target engagement. The explanation of this contradictory finding is not apparent.

The only other significant finding involves the relationship between D:CORR and G:TIME. Although significant only for the combined subsamples, the correlation for both tank commanders and gunners exhibits the same trend--namely, increased decision correctness is associated with increased gunner time. Given that D:CORR and TOT:HITS are negatively correlated and that D:CORR and D:TIME are uncorrelated, the explanation of this finding is again not readily apparent.

4.4.2.4 Hands-On Gun Laying vs Hands-On Tracking

In that they both place heavy reliance on psychomotor skill, it is more plausible to assume a relationship between the Hands-On Gun Laying (HGL) and Hands-On Tracking (HT) job samples than has been the case with the job samples previously assessed. Although the fine motor movements involved in the two tasks are not identical, there should be a negative relationship between the two measures of accuracy and between the two measures of speed. That is, short gun laying time should be associated with greater tracking distance, and less gun laying error should be associated with a greater number of tracking hits. If these relationships exist, they should be stronger for the two measures of speed than the two measures of error in that there is little concern in gun laying for the precision of the gun lay.

Additional predictions can be generated. Since tracking is not normally accomplished at the TC station, it would be plausible to expect the correlations to be higher at the gunner station. However, a counter prediction can be offered; namely, since the nature of the psychomotor movements in gun lay and tracking at the TC station bear a closer resemblance than do those at the gunner's station, there should be higher correlations at the TC station. Finally, given that tank commanders have had practice and experience at both tasks, whereas gunners have not, it would not be surprising to find the correlations higher for tank commanders than gunners.

As Table D-7 indicates, there is some statistical evidence that the HGL and HT job samples are related. Of the eight relationships involving the primary variables, two are significant for tank commanders, two for gunners, and four for the combined subsamples. The number of significant correlations exceed chance expectations for the combined subsamples ($p = .004$) and approaches significance ($p = .051$) for the tank commanders and gunners.

In accordance with expectation, the correlations between GL:TIME and tracking distance and between GL:ERROR and number of tracking hits are negative in all cases. However, contrary to prediction, the relationships are stronger for the measures of accuracy than the measures of speed. Correlations relating the measures of accuracy are significant in all cases, whereas for the mea:ures of speed, the correlations are significant only for the combined subsamples.

There is no evidence to suggest that these relationships are stronger for tank commanders or stronger at either crewstation. This latter finding suggests the operation of the two counteracting factors--the fact that tracking is not normally accomplished at the TC station and that the correspondence in motor movements between gun laying and tracking is higher at the TC station (see previous discussion).

4.4.2.5 Hands-On Gun Laying vs Hands-On Target Engagement

Although they both involve psychomotor skills, the Hands-On Gun Laying (HGL) and the Hands-On Target Engagement (HTE) tasks are different in many important ways. First, different controls are involved--the tank commander's power control handle versus the gunner's cadillac control. Secondly, the HGL task required zeroing-in on a target point with the naked eye while the HTE task afforded the use of the reticle cross hairs to zero-in on the target. Thirdly, gun laying involves large gross movements of the gun tube, whereas engagement of a target at the gunner's station involves small precision movements. Based upon the above differences, a relationship between the HGL and HTE job samples is not necessarily expected.

If relationships exist and they are plausible, one should find a positive correlation between gun laying time and gunner engagement time and a negative correlation between gun laying error and number of engagement hits. These relationships are more likely to occur for the time measures in that emphasis in both tasks is on the speed of responding, whereas precision accuracy is only the concern of one task (i.e., HTE). Furthermore, in that tank commanders have had practice and experience on both tasks, whereas gunner's have had experience only from the gunner's station, the relationships should be stronger for tank commanders than gunners.

Table D-8 (Appendix D) shows that the measures are related as expected. For both tank commanders and gunners, GL:TIME and G:TIME are positively related, whereas GL:ERROR and TOT:HITS are negatively correlated. The relationships are stronger for tank commanders than gunners in that significance was obtained for the former group and not the latter. The number of significant correlations for tank commanders (2 out of 6) exceeds chance expectations ($p = .031$). Therefore, there is some statistical evidence of a relationship between the HGL and HTE job samples but only for tank commanders.

4.4.2.6 Hands-On Tracking vs Hands-On Target Engagement

Again, due to task differences, a relationship between Hands-On Tracking (HT) and Hands-On Target Engagement (HTE) is not necessarily expected. The major difference centers around the nature of the gun tube movement. The HT task involves continuous movement of the gun tube along a snakeboard track at a subject determined pace. The HTE task, on the other hand, requires short precision movement of the gun tube to the center of mass of a stationary target at a rapid pace.

If relationships are obtained and they are plausible, then tracking distance should be negatively related to gunner engagement time and the number of tracking hits should be positively related to the number of engagement hits. Additionally, if relationships exist, they are more likely to be associated with tracking at the gunner station since tracking from the TC station is a relatively novel task.

The data (cf, Table D-9) support the contention that the two speed measures would be negatively related. In all cases, greater tracking distance is associated with shorter engagement times. These correlations are significant, however, only for gunners and for the G:TIME measure. Quite unexpectedly, these correlations are significant for tracking at both crewstations. The fact that the relationships between the two speed measures is stronger for gunners than tank commanders is plausible in that gunners have had more recent practice at both tasks. The finding of significant relationships with tracking at the TC station is also understandable in that speed of tracking at the two crewstations is significantly and positively correlated (cf, Section 4.3.6).

The measures of accuracy on the two tasks are not significantly correlated. More importantly, however, correlations between tracking hits and engagement hits are in the expected positive direction only for tank commanders. For gunners, both at the TC and gunner station, a greater number of tracking hits is associated with fewer target engagement hits. The reason for this discrepancy is not apparent.

4.4.3 Bivariate Relationships Between the Computer and Hands-On Job Samples

4.4.3.1 Computer Panel vs TC Decision Making

Despite that they are both cognitive tasks, there is no a priori reason to expect a statistical relationship between the Computer Panel (CP) and TC Decision Making (TCD) job samples. First, the cognitive skills which are involved are different. The CP job sample requires the ability to assimilate, remember, and apply new procedures, whereas the TCD job sample requires the ability to remember and apply past knowledge. Secondly, the TCD task emphasizes speed of responding, whereas the CP task does not. As a consequence, there should be little relationship between a measure of correctness constrained by time (TCD task) and a measure of correctness unconstrained by time (CP task). Finally, the measures of time in the two tasks are not parallel; completion time does not have the same performance component as does decision speed.

The results of the statistical analysis (cf, Table D-10) support the contention that there would be no statistical relationship between the CP and TCD job samples. Out of the eight relationships involving the primary measures, only one is significant for gunners and none are significant for tank commanders. Obviously, the number of significant relationships does not exceed chance expectations. The only significant correlations is that between CST:TIME and D:TIME for gunners. For gunners, longer self-test completion times are associated with longer decision times.

Despite the lack of significance, there is an interesting pattern which emerges among the correlations. For tank commanders, the two measures of correctness are negatively correlated as are the two measures of time. On the other hand, for gunners, positive correlations are obtained both between the two measures of correctness and between the two time measures. The positive correlations for gunners are perfectly understandable. The longer the decision time, the longer the time to complete the computer tasks. Correspondingly, the more time taken to perform a task, the greater the likelihood of being correct.

The negative relationships for tank commanders can be explained by making several assumptions. First, the better tank commanders (as indexed by shorter decision time) may adopt the strategy of taking a longer time on the novel computer tasks to maximmize the likelihood of being correct. Secondly, the better tank commanders in a multiple engagement situation may decide to open fire on any reasonable target rather than waste time deciding which target to engage. As a consequence, they may not necessarily engage the "correct" target as defined by the rules of engagement. Therefore, there would be tendency for the measures of correctness and the measures of time to be negatively correlated.

4.4.3.2 Computer Panel vs Hands-On Gun Laying

On an a priori basis, the CP job sample is even less likely to be related to the Hands-On Gun Laying (HGL) job sample than was the case for the TCD job sample. There are major task differences, most of which are associated with the fact that the CP job sample is primarily a cognitive task, whereas the HGL job sample is a psychomotor task. For example, the CP job sample taps procedural accuracy under conditions which do not emphasize speed of responding. On the other hand, the HGL job sample involves psychomotor accuracy under conditions where speed of responding is crucial.

Table D-11 contains the correlations between the CP and HGL measures. As can be seen, the evidence for a statistical relationship between the CP and HGL job samples is at best weak. Of the eight relationships involving the primary measures, only two are significant for tank commanders and none are significant for gunners. Although the number of significant correlations approaches significance for tank commanders ($p = .051$), one of the two significant correlations (i.e., that between ECD:TIME and GL:ERROR) involves a relationship between two variables which it does not make sense to examine on an a priori basis.

If a prediction had been advanced regarding the two time measures, one would have expected a positive relationship. Namely, shorter completion time should be associated with shorter gun laying time. Positive but non-significant correlations between completion time and gun laying time are obtained for gunners. However, for tank commanders, the relationship between the two time measures is in the opposite direction and stronger. The negative relationship between the two time measures for tank commanders achieved significance only in the case involving ECD:TIME and GL:ERROR. This negative relationship is understandable if it is assumed that the better tank commanders (as indexed by shorter gun laying times) take longer on the novel computer tasks to increase their accuracy.

The relationships between the two accuracy measures, although nonsignificant and not predicted, are plausible in that they are generally (with one exception) negative. Namely, there is a weak trend for a greater number of correct on the CP tasks to be associated with smaller gun laying error.

4.4.3.3 Computer Panel vs Hands-On Tracking

There are major task differences between the CP job sample and the Hands-On Tracking (HT) job sample. The CP job sample is a cognitive task which emphasizes procedural correctness in a situation relatively unconstrained by time pressure. On the other hand, the HT job sample is a psychomotor task which simultaneously emphasizes both speed and accuracy. In that the tasks have different performance components, there is no a priori reason to expect a relationship between the CP and HT job samples.

Inspection of Table D-11 reveals that there is no statistical evidence for a relationship between the CP and HT job samples. Significance is obtained for only one of the sixteen relationships involving the major variables, and that is only for gunners. There are no significant relationships for tank commanders. The significant correlation which is obtained for gunners involved ECD:TIME and G:HITS. Namely, for gunners, longer completion times on the Enter/Check Data task are associated with fewer tracking hits at the gunner station. Although there is no a priori reason to examine a relationship involving a measure of accuracy on one task with a measure of speed on another, this negative correlation is plausible if one assumes some general underlying gunnery skill.

Despite the low and nonsignificant correlations, a pattern emerges in the direction of the relationships. In all cases except one, the two measures of accuracy are positively correlated, with greater correctness on the CP tasks being associated with a greater number of tracking hits. The two time measures, on the other hand, are negatively related. There is a weak trend for longer completion times on the CP tasks to be associated with shorter tracking distances. This pattern in the direction of the relationships is consistent with what would have been expected had prediction been advanced.

4.4.3.4 Computer Panel vs Hands-On Target Engagement

In that the CP job sample is a cognitive task and that Hands-On Target Engagement (HTE) is basically a psychomotor task, there is no reason to expect that performance on these two job samples would be related. Although the HTE task does have cognitive elements, other task differences mitigate the likelihood of finding a relationship. The most notable of these task differences involves the differential emphasis on speed of responding. Performance under conditions which emphasize the speed of responding (HTE job sample) is not likely to be related to performance under conditions which do not emphasize speed (CP job sample).

As expected, there is no statistical evidence for a relationship between the CP and HTE job samples. Of the twelve relationships involving the primary variables, Table D-13 (Appendix D) shows that only two correlations are significant for tank commanders, and none are significant for gunners. The number of significant correlations for tank commanders does not exceed expectations based upon chance ($p = .099$). Furthermore, one of the significant correlations for tank commanders involves a relationship which one would not necessarily have examined on an a priori basis--namely, a correlation between a measure of accuracy on one task (ECD:CORR) and a measure of time on another (G:TIME). That is, for tank commanders, the number correct on the Enter/Check Data task is negatively related to gunner engagement time. This relationship is understandable if one assumes some general underlying tank gunnery skill.

In addition to the lack of significance, no consistent interpretable pattern emerges among the correlations. This is particularly true in the case of relationships involving completion time with gunner engagement time, where the direction of the correlations vary in an unpredictable fashion across CP tasks (Enter/Check Data vs Self-Test) and across crew position (tank commander vs gunner).

4.4.3.5 Computer Tracking vs TC Decision Making

In that the CT task involves primarily psychomotor skill and the TCD task involves cognitive skill, there is no a priori reason to expect a relationship between these two job samples. However, given the novelty of the Computer Tracking task (e.g., unfamiliar controls and control dynamics, lack of a background against which to assess target movement), there is probably a heavier cognitive component than would normally be the case in that the tank crewmen must constantly think about their control movements. On this basis, significant relationships between the CT and TCD job samples are possible but unlikely.

Inspection of Table D-14 shows that there is no statistical evidence for a relationship between the CT and TCD job samples. Significance is obtained for only one of the twelve relationships involving the primary variables, and that is for the combined subsamples. None of the correlations for tank commanders or gunners achieve significance. The only significant correlation involved a cross comparison--a measure of accuracy on one task with a speed measure on another task. Namely, for tank commanders, the greater the tracking error on the easy task, the greater the decision time.

In examining the pattern of intercorrelations as to its plausibility, comparisons were restricted to only the accuracy measure in that there were no speed measures for the CT task. Under the assumption that there is some general task gunnery aptitude, it would be expected that there would be a positive correlation between time-on-target and decision correctness and a negative correlation between tracking error and decision correctness. The expected pattern occurred for gunners (with only one exception) but not for tank commanders. For tank commanders, tracking error and decision correctness are negatively related in two of three cases (as expected). However, contrary to expectation, time-on-target and decision correctness are negatively related.

4.4.3.6 Computer Tracking vs Hands-On Gun Laying

Since Computer Tracking (CT) and Hands-On Gun Laying (HGL) both place heavy reliance on psychomotor skills, there is potential for a possible relationship. However, there are important task differences. The HGL task places emphasis on the speed at which the gun is layed, whereas there is no element of speed in the CT task. The HGL job sample requires tank crewmen to bring the gun tube toward the target dot in one rapid movement. The CT job sample, on the other hand, requires tank crewman to move the target to the reticle and maintain the target-reticle overlap over a period of time. Finally, in the CT job sample, accuracy is achieved with the aid of a reticle, whereas in the gun laying task, accuracy is accomplished with the aid of the naked eye. Based upon the above task differences, it is reasonable to expect that if relationships exist between the CT and HGL job samples, they are not likely to be strong.

Given that there is no measure of speed for the CT job sample, predictions can be legitimately made only between the accuracy measures on the two tasks. In advance of collecting the data, it would be reasonable to expect that gun laying error would be negatively related to time-on-target and positively related to tracking error.

However, there are other factors which complicate the predictions regarding the accuracy measures. First, speed and not precision accuracy, is the concern in gun laying. Therefore, the degree of accuracy in gun laying may be uncorrelated with the degree of tracking accuracy. Second, the differential experience of tank commanders and gunners in gun laying makes prediction difficult. On the one hand, given that tank commanders have had experience at both tasks, whereas gunners have not, it is plausible to expect stronger relationshps for tank commanders than gunners. On the other hand, the greater experience of tank commanders make them less prone to worry about the accuracy of their gun lay. As a consequence, there may be little correspondence between the measures of accuracy on the two tasks for tank commanders. Given their lack of experience in gun laying, the gunners may place a greater emphasis on accuracy; therefore, relationships between measures of accuracy would be more likely.

Table D-15 (Appendix D) presents the intercorrelations between the CT and HGL measures. As Table D-15 indicates, there is some evidence for a statistical relationship between the CT and HGL measures, but only for gunners. Of the twelve relationships involving the primary variables, four are significant for gunners and none are significant for tank commanders. Although the number of significant correlations for gunners is above chance level (p = .002), only two of them involved relationships between tracking accuracy and gun laying accuracy.

The pattern of correlations between measures of accuracy on the two tasks is as predicted for gunners. Namely, small gun laying error is associated with greater time-on-target and small tracking error. These relationships are significant, however, only in the case of EASY:TOT and HARD:ERROR.

For tank commanders, the direction of the relationship between the measures of accuracy on the two tasks are in accordance with prediction only in the case of the easy and hard tracking. For the moderate tracking tasks, however, gun laying error was positively correlated with time-on-target and negatively correlated with tracking error.

4.4.3.7 Computer Tracking vs Hands-On Tracking

In that the CT and the Hands-On Tracking (HT) job samples were designed to tap the same behavior at two different levels of abstraction, there is an a priori basis for expecting a relationship between the two job samples. More specifically, it was predicted that the number of tracking hits (HT) would be positively related to time-on-target (CT) and negatively related to tracking error (CT). Additionally, in that gunners have had more recent practice in tracking, it was expected that these relationships would be stronger for gunners than tank commanders. Finally, since tracking is a relatively novel task at the TC station, it was anticipated that the above relationships would be more likely to occur at the gunner's staion. Prediction regarding the tracking distance measure (HT task) cannot be advanced since there is no corresponding speed measure in the CT task.

Given important task differences, the relationships between the accuracy measures on the two tasks are not expected to be strong. For example, the computer joystick control and its control dynamics differ from either those of the tank commander's power control handle or the gunner's cadillac controls. The CT task involves tracking in absence of background cues, whereas the HT task involves tracking with background cues present. Finally, accuracy of tracking in the HT task is governed by the speed of tracking; but in the CT task, speed is an irrelevant factor.

The intercorrelations between the performance measures on the CT and HT job samples are tabled in Appendix D, Table D-14. Inspection of Table D-14 shows that four of the 24 correlations involving the primary measures are significant for gunners, whereas only one is significant for tank commanders. The number of significant correlations for gunners exceeds chance expectations (p = .024). Interestingly, three of the four significant correlations for gunners involve the tracking hits measure at the TC station. Therefore, there is some statistical evidence for a relationship between the CT and HT job samples for gunners but primarily at the TC station.

The pattern of correlation between the measures of accuracy on the two tasks for gunners exactly conforms to expectations with only one exception--

the relationships are much stronger at the TC station. For the gunners, a greater number of tracking hits on the HT task is associated with a longer time-on-target and less tracking error on each of the CT tasks. The fact that these relationships are stronger at the TC station is understandable if one assumes that behavior on two novel tasks are more likely to be correlated than is behavior between a novel and a well practiced task. Additionally, the computer joystick and the TC power control handle are more similar than the joystick and gunner's cadillac.

For tank commanders, the correlations between measures of accuracy on the two tasks fit no predictable pattern. The correlations for the same two measures are sometimes positive and sometimes negative. The direction of the correlations vary in an inconsistent fashion across CT tracking tasks, CT tracking measures, and across HT crewstations.

4.4.3.8 Computer Tracking vs Hands-On Target Engagement

Although they both involve psychomotor skill, the CT and Hands-On Target Engagement (HTE) job samples are not necessarily expected to be related because of essential task differences. In the HTE job sample, there is an emphasis in the speed of responding, whereas in the CT job sample, there is no element of speed. The HTE task involves quickly moving the center of the reticle to the center of mass of a stationary object. The CP task, on the other hand, requires bringing a continously moving object (target) to the center of the reticle and maintaining the target reticle overlap over a period of time.

If relationships exist and they are plausible, then the number of target engagement hits should be positively related to time-on-target and negatively related to tracking error. Again, since there are no measures of speed on the CT task, predictions must be limited to the measures of accuracy.

Inspection of Table D-17 reveals that there is good statistical evidence for a relationship between the CT and HTE job samples, but only for gunners. Out of the eighteen relationships involving the primary variables, eight correlations are significant for gunners, whereas only two are significant for tank commanders. However, the number of significant correlations is above chance level only for gunners ($p = .000$).

It is noteworthy that six of the significant correlations (five for gunners and one for tank commanders) involved relationships between the accuracy measures. However, Table D-17 shows that the pattern of relationships between the measures of accuracy in the two tasks is exactly opposite to prediction in all cases for both tank commanders and gunners; that is, a greater number of target engagement hits is associated with a shorter time-on-target and a larger tracking error. The reason for this unexpected finding is not apparent.

Although there is no a priori reason to make a comparison between a measure of speed on one task and a measure of accuracy on another, a consistent pattern emerges between target engagement time and the tracking accuracy measures. The target engagement time measures (TC:TIME, G:TIME) are negatively correlated with time-on-target and positively correlated with tracking error (with two exceptions). In support of this pattern, three of the significant correlations for gunners involved relationships between TC:TIME and the tracking accuracy measure. The reason that these relationships are stronger for the tank commander time segment is not clear.

4.4.3.9 Computer Target Engagement vs TC Decision Making

Although the Computer Tracking Engagement (CTE) job sample involves psychomotor skill, it has major components of decision making. Unlike Hands-On Target Engagement, the tank crewman in the CTE task must decide when to switch from 3X to 10X, when to lase, and whether or not to relase given the presence or absence of the multiple return bar. Although the time in the 3X and Laser 1 segments is largely determined by the time to acquire the target, the tank crewmen must decide when to press the appropriate button. On the other hand, the decision as to whether or not to relase is mostly a cognitive function. The time to make the relase decision is obviously going to affect total time in 10X. Therefore, given that the CTE task has a heavy decision making component, it is not unreasonable to expect a relationship between the TCD and CTE job samples.

If relationships exist, it is plausible that they are most likely to occur between decision time on the TCD task and the CTE time measures and between number of correct decisions on the TCD task and procedural errors on the CTE task. More specifically, decision time is expected to be positively related to time in 3X and time in 10X, whereas decision correctness is predicted to be negatively related to procedural errors. There is no reason to predict a relationship betweeen decision correctness and either 3X or 10X distance error in that the former is a measure of cognitive accuracy and the latter are measures of psychomotor accuracy.

Table D-18 contains the correlations between the CTE and TCD performance measures. As Table D-18 indicates, two of the 10 relationships involving the primary variables are significant for tank commanders. However, none are significant for gunners. The number of significant correlations for tank commanders approaches being above chance level (p = .075). In that the two significant correlations occur in cells in which a relationships is expected, they probably do not represent a chance occurrence.

As expected, the correlation between D:TIME and 3X:TIME and between D:TIME and 10X:TIME are all positive. Shorter decision times on the TCD task are associated with shorter times in the 3X and 10X segments on the HTE task. These correlations are significant, however, only for tank commanders.

The correlations beteen PROC:ERROR and D:CORR are in the expected direction. Namely, the greater the number of procedural errors on the HTE task, the fewer the number of correct decisions on the TCD task. Despite being in the expected direction, this correlation was not significant for either tank commanders or gunners.

4.4.3.10 Computer Target Engagement vs Hands-On Gun Laying

In that they both involve psychomotor skill, it is not unreasonable to expect a relationship between the CTE and HGL tasks. However, due to the extra cognitive demands of the CTE task, the correspondence between measures on the two tasks are likely to be low.

There are other task differences which would tend to lessen the relationship between the CTE and HGL job samples. For example, the two tasks utilize different controls with different dynamics--the computer joystick for the CTE task versus the tank commander's power control handle for the HGL task. Moreover, the HGL task involves achieving accuracy with the use of the naked eye, whereas the CTE task involves the use of the cross hairs of the reticle to achieve accuracy.

If plausible relationships exist, there is likely to be a positive correlation between the time measures on the two tasks as well as a positive correlation between the accuracy measures (excluding procedural errors) on the two tasks. Table D-19 (Appendix D) reveals that the direction of the relationships are as expected (with only one exception). In general, larger gun laying error is associated with larger distance error in both the 3X and 10X time segments. Although the relationships tend to be weaker, longer gun laying times correspond with longer time spent in the 3X and 10X segments.

Despite the fact that the correlations are in the expected direction, there is little statistical evidence to indicate a relationship between the CTE and HGL job samples. For both tank commanders and gunners, only one of the ten correlations involving the major variables achieved significance. In both cases, 3X:ERROR(AVG) was significantly related to GL:ERROR. Nevertheless, the number of significant correlations does not exceed expectations based upon chance ($p = .315$).

4.4.3.11 Computer Target Engagement vs Hands-On Tracking

Since the CTE and HT job samples both involve psychomotor skill, it is not implausible to expect a relationship between their performance measures. However, if relationships exist, they are not likely to be strong because of major task differences. For example, the CTE task has a heavy cognitive component in that new procedures must be learned and applied, whereas the HT task relies primarily on psychomotor skill, In addition, the two tasks have different psychomotor components. The CTE task requires bringing the reticle over a stationary target in one quick rapid movement, whereas the HT task requires continuous movement of the gun tube at a subject-determined "pace" to stay within limits of the width of the track. Lastly, the different tasks involve different controls and control dynamics.

If plausible relationships exist, it is reasonable to expect a negative correlation between distance error in the CTE task and number of tracking hits and a positive correlation between the engagement time measures and the distance tracked. In that gunners have had more recent practice in tracking from the gunner's station and that the CTE task is basically a gunner task, these relationships are more likely to be stronger for gunners in the gunner station. Predictions regarding relationships with tracking at the TC station are difficult to advance because of the novelty of the task (i.e., tracking is not normally accomplished from the TC station).

The interrelationships which actually emerged are shown in Table D-20. As can be seen, an inconsistent pattern resulted. For the 3X time segment, three of the four relationships between 3X:ERROR(AVG) and number of tracking hits are negative as predicted. However, for the 10X time segment, correlations between number of tracking hits and 10X:ERROR(AVG) are negative for tank commanders, but positive for gunners. When considering the relationships between the speed measures (3X:TIME and 10X:TIME vs tracking distance), the correlations are all positive for gunners (as expected) but are negative in three out of the four cases for tank commanders. More importantly, only one of the presumed relationships achieved significance--that between 10X:TIME and G:DISTANCE for gunners; the longer the time in 10X, the greater the tracking distance.

Although the expected pattern of correlations between the two measures of accuracy and between the two measures of speed did not emerge, there is some statistical evidence that the CTE and HT job samples are related, but for gunners only. Out of the 20 relationships involving the primary measures, five correlations are significant for gunners, whereas only one is significant for tank commanders. The number of significant correlations for gunners exceeds chance expectation ($p = .002$).

Four of the five significant correlations for gunners involve a cross comparison--that is, a measure of accuracy in one task with a measure of speed on another. For gunners, the greater the time in either 3X or 10X, the fewer the number of tracking hits. This relationship was significant at both the TC and gunner's station. Although not predictable in advance, these relationships are understandable if one assumes some general underlying tank gunnery aptitude.

4.4.3.12 Computer Target Engagement vs Hands-On Target Engagement

Relationships between the CTE and the HTE job samples are definitely expected in that the CTE task was designed to tap the same behavior as the HTE task but at a different level of abstraction. Although the nature of the accuracy measures are different, there should be a negative correlation between target distance error on the CTE task and the number of target engagement hits on the HTE task. Furthermore, time in 3X and time in 10X on the CTE task should be positively correlated with gunner engagement time on the HTE task.

The relationships which emerge, however, are not likely to be strong due to important task differences. Aside from the obvious equipment differences, there are differences in the cognitive aspects of the two tasks. In a CTE task, the tank crewman must detect and locate the target, whereas in the HTE test, the confederate occupying the tank commander station accomplishes this task. The CTE task involves learning and applying a new set of procedures (those associated with lasing) to the target engagement sequence. In the CTE task, the tank crewman must make decisions as to when and when not to lase. None of these cognitive aspects are present in the HTE task.

Contrary to expectation, there is no evidence for a statistical relationship between the CTE and HTE job samples. Inspection of Table D-12 reveals that significance is obtained in only one of fifteen cells involving relationships between the primary variables and that relationship involves a cross comparison. For gunners, there is a significant positive correlation between 3X:TIME and TOT:HITS.

In addition to the failure to obtain significance, the pattern of interrelationships is not completely in accordance with prediction. In three out of the four cases, the relationship between target distance error on the CTE task [(3X:ERROR(AVG), 10X:ERROR(AVG)] and number of hits on the HTE task is negative as expected. However, contrary to expectation, the correlation between the time measures on the CTE task (3X:TIME and 10X:TIME) and gunner engagement time in the HTE task is negative in all cases. The reason for this latter finding is not clear.

4.4.4 Canonical Relationships Between Job Samples

Despite the lack of statistical evidence using bivariate correlation, there is a consistent pattern between the measures of accuracy and between the measures of time (speed) for many of the job sample pairings. To investigate significance in the pattern of relationships between sets of job sample measures, canonical correlations were computed for each of the 21 job sample pairings.

Canonical correlation is a technique which measures the degree of relationship between two sets of variables. In that canonical correlation simultaneously considers sets of variables, it is possible to detect significance in the pattern of relationships between two job samples--something that might go undetected with bivariate correlation. In addition, through canonical correlational analysis, it is possible to assess the degree of variability which two sets of measures share in common. As with bivariate correlation, the canonical correlation coefficient can be squared to determine the proportion of shared variance.

Table 4-17 contains the canonical correlations which resulted for each of the 21 job sample pairings. Before elaborating upon the results, three observations are noteworthy. First, the magnitude of the canonical correlation coefficient is very much affected by the number of variables in the two variable sets. The greater the number of variables included in the two variable sets, the higher the canonical correlation and the greater the proportion of variability shared in common. Therefore, given that each job sample pairing contains a different number of measures in the two variable sets, the magnitude of the canonical correlations cannot be compared across job sample pairings. Second, as with all multivariate techniques, the power of the statistical test is affected by both the number of variables and the sample size. As the number of variables increase, the sample size must be increased to achieve the same statistical power. Given the number

TABLE 4-17. CANONICAL CORRELATIONS BETWEEN MEASURES OF PERFORMANCE ON 7 JOB SAMPLES FOR TANK COMMANDERS (TC), GUNNERS (G), AND COMBINED (TCG) SUBSAMPLES

Job Sample	Variables in Set	Sample	CT	CTE	TCD	HGL	HT	HTE
Computer Panel (CP)	ECD:CORR	TC	.894	.638	.476	.804***	.364	.833***
	ECD:TIME	G	.689	.521	.535*	.393	.519	.461
	CST:CORR	TCG	.423	.464	.316	.286	.383	.434
	CST:TIME							
Computer Tracking (CT)	EASY:TOT	TC		.690	.568	.692	.735	.839
	EASY:ERROR	G		.635	.422	.574**	.577**	.585
	MOD:TOT	TCG		.536	.316	.420	.436	.477
	MOD:ERROR							
	HARD:TOT							
	HARD:ERROR							
Computer Target Engagement (CTE)	PROC:ERROR	TC			.599	.561	.579	.543
	3X:ERROR (AVG)	G			.399	.539**	.544	.392
	3X:TIME	TCG			.406	.480***	.407	.368
	10X:ERROR (AVG)							
	10X:TIME							
TC Decision Making (TCD)	D:CORR	TC				.387	.451	.415
	D:TIME	G				.322	.399	.427
		TCG				.302*	.332*	.400*
Hands-On Gun Laying (HGL)	GL:TIME	TC					.540	.482
	GL:ERROR	G					.602***	.336
		TCG					.584****	.346**
Hands-On Tracking (HT)	TC:HITS	TC						.397
	TC:DIST	G						.555
	G:HITS	TCG						.393
	G:DIST							
Hands-On Target Engagement (HTE)	TOT:HITS	TC						
	TC:TIME	G						
	G:TIME	TCG						

* p ≤ .10
** p ≤ .05
***p ≤ .01
****p ≤ .001

of variables included in most of the job sample pairings, the sample sizes for tank commanders and gunners are not large enough to achieve sufficient statistical power. Therefore, the .10 significance level was adopted in assessing the significance of the canonical correlations. Finally, in that the power of the multivariate statistical test is affected by the number of variables, care must be taken in interpreting the statistical results. It is possible for a large canonical correlation based upon several variables to be nonsignificant, whereas a smaller canonical correlation based upon fewer variables may be significant.

Despite the fact that 43 percent of the bivariate correlations are less than 0.100, the canonical correlations indicate a much higher degree of relationship between the samples. For tank commanders, the canonical correlations range from .364 to .894, whereas for gunners, the canonical correlations range from 0.322 to 0.689. If these values are then squared, the proportion of shared variance between sets of measure for the job sample pairings vary from 13 to 89 percent for tank commanders and from 10 to 47 percent for gunners.

Only two of the canonical correlations achieve significance for tank commanders--that between the Computer Panel and Hands-On Gun Laying job samples and that between the Computer Panel and Hands-On Target Engagement job samples. Upon inspection of the bivariate correlations involving these two job sample pairings (cf, Tables D-11 and D-13), the following facts emerge for the CP-HGL pairing. The correlations between the two time measures are opposite to what one might expect--namely, short gun laying times are associated with long CP completion times. Furthermore, most of the CP-HGL pattern correspondence is due to accuracy-time relationships. As CP completion time increases, there is a corresponding increase in gun laying error. Conversely, there is also a tendency for short gun laying times to be associated with a greater number correct on the CP task.

In the CP-HTE comparison, the correlations involving the two accuracy measures are positive as expected. However, there is only partial confirmation of the fact that the two time measures would be positively correlated as well. As might be expected, the correlations of the number correct on the CP task with gunner engagement time are both negative.

Five canonical correlations are significant for gunners. In four of the five comparisons (CP vs TCD, CT vs HGL, CT vs HT, HGL vs HT), the correlations between the two measures of accuracy and the two measures of time are exactly in accordance with expectation. The CT-HT comparison is noteworthy in that the two tasks were specifically designed to tap the same behavior but at two different levels of abstraction.

For the fifth significant gunner comparison (CTE vs HGL), the results are not in complete accordance with what might be expected. The relationships were strong and in the expected direction only for the 3X time segment. For the 3X segment, 3X:ERROR and GL:ERROR are positively correlated, as are 3X:TIME and GL:TIME. However, for the 10X segment, only the prediction regarding the time relationships is confirmed.

Six canonical correlations are significant for the combined subsamples. In that use of separate subsamples may not permit a powerful enough statistical test of the relationships for tank commanders and gunners, the significance of the combined subsamples may be used in cases where the relationships are the same for tank commanders and gunners.

Five out of the six significant overall canonicals deal with relationships among the hands-on job samples. For two of these hands-on comparisons (HGL vs HT and HGL vs HTE), the accuracy and speed measures are correlated exactly as expected for both tank commanders and gunners. For two other hands-on comparisons (TCD vs HGL and TCD vs HTE), there is complete agreement between tank commanders and gunners in showing a positive relationship between the two time measures. That is, short decision times were associated with short gun laying and gunner engagement times. When considering the accuracy/correctness relationships, tank commanders and gunners both exhibit a trend toward greater decision correctness (TCD) being associated with a fewer number of engagement hits (HTE). Obviously, this latter trend is opposite to expectation. For the last significant overall hands-on comparison (TCD vs HT), both tank commanders and gunners exhibit the expected positive relationship between decision correctness and number of tracking hits in all cases. However, decision time and tracking distance are negatively related in only three out of the four cases.

The last significant overall canonical relationship involves comparison of the CTE and HGL job samples. In that comparison, short gun laying times are always associated with shorter times in 3X and 10X. In addition, as expected, 3X:ERROR and GL:ERROR are positively correlated for both tank commanders and gunners.

4.4.5 Summary of Interrelationships Among Job Samples

The purpose in examining the interrelationships among job samples was twofold--to establish the degree of intercorrelation among job sample measures and to determine the plausibility of the relationships which emerged (construct validity). The data related to each purpose is summarized below.

4.4.5.1 Intercorrelation Among Individual Measures of Job Sample Performance

Table 4-18 summarizes the degree of intercorrelation which exists among the predictor measures. As can be seen, the bivariate correlations are in general low, with approximately one-half of the correlations being 0.150 or below. Correspondingly, the percentage of bivariate correlations that are significant are also low. In only a few cases do the number of significant correlations exceed chance expectations.

The low degree of intercorrelation among the job sample measures is ideal from a measurement perspective. With a low intercorrelation among the job sample measures, there is a greater probability of finding measures which will add to the ability to predict the criteria variables in Phase II of the present investigation.

TABLE 4-18. SUMMARY OF THE DEGREE OF RELATIONSHIP AMONG JOB SAMPLE MEASURES

Job Sample Comparison	Tank Commanders			Gunners		
	Median Bivariate Correlation	Percentage of Bivariate Correlations Which are Significant	Bivariate Correlations Above Chance Level	Median Bivariate Correlation	Percentage of Bivariate Correlations Which Are Significant	Bivariate Correlations Above Chance Level
CP vs CT	.152	0(24)[1]	No	.062	0(24)	No
CP vs CTE	.149	5(20)	No	.235	10(20)	No
CT vs CTE	.118	3(30)	No	.091	20(30)	Yes
TCD vs HGL	.054	25(4)	No	.181	0(4)	No
TCD vs HT	.067	25(8)	Yes	.075	25(8)	Yes
TCD vs HTE	.222	0(6)	No	.069	0(6)	No
HGL vs HT	.070	25(8)	Yes	.170	25(8)	Yes
HGL vs HTE	.164	33(6)	Yes	.120	0(6)	No
HT vs HTE	.153	0(13)	No	.197	17(12)	No
CP vs TCD	.065	0(18)	No	.131	6(18)	No
CP vs HGL	.230	25(8)	Yes	.113	0(8)	No
CP vs HT	.057	6(16)	No	.088	6(16)	No
CP vs HTE	.117	17(12)	No	.151	0(12)	No
CT vs TCD	.150	0(12)	No	.078	0(12)	No
CT vs HGL	.166	0(12)	No	.239	33(12)	Yes
CT vs HT	.113	4(24)	No	.190	17(24)	Yes
CT vs HTE	.197	11(18)	No	.268	44(18)	Yes
CTE vs TCD	.245	20(10)	No	.123	0(10)	No
CTE vs HGL	.075	10(10)	No	.102	10(10)	No
CTE vs HT	.077	5(20)	No	.115	20(20)	Yes
CTE vs HTE	.126	0(15)	No	.093	7(15)	No

[1] Number in parentheses indicates number of correlations upon which percentage is based.

4.4.5.2 Construct Validity of the Job Samples

To assess construct validity, logical deductions were first made about the direction of relationships between job samples and job sample measures. These were then compared to the obtained correlations to determine if the presumed relationships occurred. To evaluate the criterion related validity of the job samples, job sample measures were correlated with measures of past success at Annual Qualifications.

If a single tank gunnery aptitude underlies job sample performance, then there should be a high intercorrelation among job sample measures, both between and within job samples. In that there were relatively high intercorrelations among the performance measures within job samples but not between job samples, the evidence indicates that the job samples tap different behaviors and possibly different aptitudes.

As summarized in Table 4-18, the intercorrelation between measures of different job samples were low and with only a few exceptions were nonsignificant. In that there are major task differences between job samples, the low intercorrelations between measures of different job samples were not unexpected.

However, despite the lack of significance, a predictable pattern in the direction of relationships emerged among the job sample measures which cannot be attributed solely to random chance variation. That is, based upon knowledge of tank gunnery and human performance, predictions about the direction of relationships were advanced. In many cases the presumed direction of relationship was confirmed by the data. Table 4-19 summarizes those job sample comparisons in which the predicted direction of relationships emerged.

Table 4-19 indicates whether or not the presumed direction of relationship was confirmed by the data for each of the 21 job sample comparisons and separately for accuracy/correctness and speed/time relationships. In constructing the table, a "Yes" was placed in a cell if and only if two-thirds of the obtained relationships were in the expected direction and a consistent relationship pattern emerged. Table 4-19 also indicates whether or not the canonical correlation was significant. In places where there was canonical significance for the combined subsamples but not in an individual tank crewman subsample, the word "Overall" appears. In that the overall canonical is a more powerful statistical test, it may be used to indicate significance when the same pattern of relationships exists for tank commander and gunners.

Inspection of Table 4-19 reveals that the expected direction of relationship occurred in 20 of the 36 cells for tank commanders and in 24 of the 36 cells for gunners. The number of cells in which the expected relationship emerged exceeded chance expectation for gunners ($p = .018$) but only approached significance for tank commanders ($p = .106$). However, when the actual percentage of correlations which are in the expected direction is used as the basis for the statistical test, the frequency of expected relationships is above chance level for tank commanders as well.

TABLE 4-19. SUMMARY OF EVIDENCE FOR CONSTRUCT VALIDITY OF JOB SAMPLES

Job Sample	Tank Commanders			Gunners		
	Direction of Relationships			Direction of Relationships		
	Accuracy/ Correctness Comparisons	Speed/ Time Comparisons	Canonical Significance	Accuracy/ Correctness Comparisons	Speed/ Time Comparisons	Canonical Significance
CP vs CT	Yes	--[1]	No	No	--	No
CP vs CTE	No	No	No	Yes	Yes	No
CT vs CTE	Yes	--	No	No	--	No
TCD vs HGL	No	Yes	Overall	Yes	Yes	Overall
TCD vs HT	Yes	No	Overall	Yes	Yes	Overall
TCD vs HTE	No	Yes	Overall	No	Yes	Overall
HGL vs HT	Yes	Yes	Overall	Yes	Yes	Yes
HGL vs HTE	Yes	Yes	Overall	Yes	Yes	Overall
HT vs HTE	Yes	Yes	No	No	Yes	No
CP vs TCD	No	No	No	Yes	Yes	Yes
CP vs HGL	No	No	Yes	Yes	Yes	No
CP vs HT	Yes	Yes	No	Yes	Yes	No
CP vs HTE	Yes	No	Yes	No	No	No
CT vs TCD	No	--	No	Yes	--	No
CT vs HGL	Yes	--	No	Yes	--	Yes
CT vs HT	No	--	No	Yes	--	Yes
CT vs HTE	No	--	No	No	--	No
CTE vs TCD	No	Yes	No	No	Yes	No
CTE vs HGL	Yes	Yes	Overall	No	Yes	Yes
CTE vs HT	Yes	Yes	No	No	No	No
CTE vs HTE	No	No	No	Yes	No	No

[1]Dash indicates that impossible to make comparison because CT job sample has no Speed/Time variables.

As Table 4-19 indicates, not only do the expected relationships occur more often for gunners, they occur more frequently for comparison involving hands-on job samples. More importantly, however, the relationships between the computer tracking and hands-on tracking and between the computer target engagement and hands-on target engagement job sample must be noted. These job samples were designed to tap the same behavior but at a different level of abstraction and, therefore, should be statistically related. For the two tracking tasks, the relationships conformed to expectation and were significant only for gunners. For tank commanders, the relationships were generally opposite to expectation and nonsignificant. For the two target engagement job samples, there was absolutely no evidence for a statistical relationship. Moreover, in three out of the four cells (cf, Table 4-19), the direction of the relationships were opposite to expectation.

To recapitulate, there was generally good evidence for the construct validity of the job samples. The job samples were designed to tap different behaviors and that they do was confirmed by the low intercorrelation among measures of different job samples. Despite the low and nonsignificant correlations, the predicted pattern of relationships emerged in the majority of cases. For two job samples in which relationships were definitely expected (i.e., computer tracking and hands-on tracking), the predicted pattern emerged and significance was obtained, but for gunners only. The major piece of disconfirming evidence came from the fact that there was no evidence for a relationship between the two target engagement tasks.

4.5 ANALYSIS OF PAST SUCCESS AT ANNUAL QUALIFICATIONS

In absence of any postpredictor criteria data, an attempt was made to assess the criterion-related validity of the job samples. Self-reported success at past Annual Qualifications was used as the criterion measure. To create a numerical score, a "1" was assigned if the tank crewman was a member of a crew which was unqualified; a "2" if he was a member of a crew which was qualified; and a "3" if he belonged to a crew which was distinguished. Thus, the higher the number, the greater the success at Annual Qualifications.

Six different success measures were generated--three associated with average success over the past eight years and three associated with success at the most recent (1981) Annual Qualifications. For each measure, data were only included if the tank crewman had been at Annual Qualifications in either the tank commander or gunner station. The six measures created were: qualification average at the TC station (QAVG:TC), qualification average at the gunner station (QAVG:G), qualification average based upon both the TC and gunner stations (QAVG:TCG), most recent qualification score if at TC station (MRQ:TC), most recent qualification score if at gunner station (MRQ:G), and most recent qualification score at either the TC or gunner station (MRQ:TCD). Since the ultimate interest is in predicting success separately for tank commanders and gunners, the criterion variables of major interest are QAVG:TC, QAVG:G, MRQ:TC, and MRQ:G.

The analysis of past success at Annual Qualifications proceeded at three levels. First, bivariate correlations were examined to determine if individual job sample measures were significantly related to past success. Second, multiple correlations were calculated to see if any job sample as a whole could predict past success. Last, to obtain some indication of the validity of the approach to be utilized in Phase II, forward stepped multiple regression analyses were performed incorporating variables from all job samples. The results of each of these analyses are presented below.

Prior to consideration of the results, several cautions must be noted regarding interpretation of the Annual Qualifications data. First, unlike number of first round hits and opening time, the qualification score described above is an insensitive metric. That is, the qualification score is a gross, global measure which is not sensitive to individual variation in performance. Add to this the fact that the qualification score in a crew based measure (i.e., based upon the crew's performance and not an individual's performance), it is not too likely that many significant relationships would be obtained.

Second, the variability in the qualification scores is very small. In most cases, the level of reported proficiency was that of "qualified." There were very few cases of "unqualified" and "distinguished" at the gunner and tank commander crew positions. This was particularly true for the measures dealing with the most recent qualification. Therefore, the variability which is being explained is small and represents that of a small select group of tank crewmen.

Third, there are two problems associated with sample size. On the one hand, there is the lack of statistical power which results from small sample sizes. The relatively small sample sizes reduce the likelihood of finding significant results. This is more of a problem for the measures based upon the most recent Annual Qualifications where only 30 tank crewmen had participated as tank commanders and 24 as gunners.

More importantly, however, the small sample sizes create an interpretative problem. When correlating the job sample measures with past success at Annual Qualifications, only valid cases in which there is a complete set of data can be employed. That is, only those subjects for which there is a job sample score and a success score can be utilized in the data analysis. Since the measures of past success are always based upon a smaller number of subjects than are the job sample measures, the job sample data for many of the subjects never enter the data analysis. In that data from a select subset of subjects is used, the relationships observed may be distorted because they may not be representative of the data for the entire population of subjects. Thus, the relationships which emerge may not be what they would have been had data for past success at Annual Qualifications been available for all subjects.

The possible problem of relationship distortion is more of a concern for the multiple correlation and multiple regression analyses than it is for bivariate correlational analyses. With the multivariate approaches, each case for data analysis is based upon more than just two variables. As a consequence, a valid case must contain data on all the variables. Therefore, cases which were valid for bivariate correlational analysis may be excluded for multiple correlation and regression analyses because there is missing data on one or more of the other variables. In addition, the multivariate approaches take into consideration the intercorrelation among the predictors. Since the number of valid cases is far less than the number of cases upon which each job sample is based, the interrecorrelations among the predictors (job samples) which result are not those reported in previous sections. Thus, the job sample measures which significantly predict past success may bear no relationship to the job sample measures which predict future success because the intercorrelational structure among the predictors is not the same.

Given the above problems, the results dealing with past success at Annual Qualifications are not considered in great depth. Rather they are examined merely to ascertain if there is any evidence for a possible relationship between the job sample measures and measures of success and to determine if the multiple regression approach of simultaneously considering all job sample measures for possible inclusion looks promising.

4.5.1 Bivariate Relationships with Past Success

The bivariate correlations with measures of past success at Annual Qualifications are presented in Appendix E, Tables E-1 through E-8. In examining the relationships which emerged, attention was focused on the combined subsample data (TCD) for four of the measures of past success--QAVG:TC, QAVG:G, MRQ:TC, and MRQ:G. The combined subsample data was utilized because the current crew position does not enter into where the crewman qualified in the past.

Close inspection of the tables reveals that the correlations are in general very low, are in many cases opposite to the direction expected, vary with measure of past success, and are significant in only six cases. Interestingly, five of the six significant correlations represent cases where the direction of the relationship is opposite to expectation.

Table 4-20 summarizes the bivariate relationships between measures of past success and army experience and the job sample measures. The table presents the percentage of bivariate correlations which are in the expected direction. Where appropriate, separate percentages are reported for the accuracy/correctness (A/C) and the speed/time (S/T) measures. In making up this table, it was assumed that it is highly desirable for tank crewmen to respond as accurately and quickly as possible. Therefore, it was expected that measures of accuracy and correctness should be positively correlated with past success, whereas measures of response time and completion time should be negatively correlated with past success. With regard to army experience, a positive relationship was assumed to represent the expected direction.

None of the correlations relating Army experience to measures of past success are significant. Moreover, the direction of the correlations with army experience follows no consistent pattern--the direction being highly

TABLE 4-20. SUMMARY OF BIVARIATE RELATIONSHIPS WITH MEASURES OF PAST SUCCESS AT ANNUAL QUALIFICATIONS

Experience/ Job Sample	Measure Dimension	Percentage of Bivariate Correlations in Expected Direction			
		Qualification at TC Station		Qualification at Gunner Station	
		QAVG:TC	MRQ:TC	QAVG:G	MRQ:G
Army Experience	--	60(5)[1]	100(5)	40(5)	0(5)
Computer Panel	A/C	0(2)	100(2)	50(2)	100(2)
	S/T	50(2)	50(2)	100(2)**	50(2)
Computer Tracking	A/C	67(6)	67(6)	0(6)	0(6)*
Computer Target Engagement	A/C	0(3)	33(3)	67(3)	100(3)
	S/T	100(2)	100(2)	100(2)	50(2)
TC Decision Making	A/C	0(1)*	0(1)*	0(1)	100(1)
	S/T	100(1)	100(1)	100(1)	100(1)
Hands-On Gun Laying	A/C	100(1)	100(1)	100(1)	0(1)
	S/T	0(1)	100(1)	0(1)	0(1)*
Hands-On Tracking	A/C	0(2)	0(2)	0(2)	100(2)
	S/T	100(2)	100(2)	0(2)	0(2)
Hands-On Target Engagement	A/C	0(1)	0(1)*	0(1)	100(1)
	S/T	100(1)	100(1)	0(1)	0(1)

[1]Number in parentheses indicates number of correlations upon which percentage is based.

*Indicates a correlation is significant in the direction opposite to expectation.

**Indicates a correlation is significant in the direction consistent with expectation.

variable across both success and experience measures. For example, all of the experience measures are positively related to success at the most recent Annual Qualifications at the TC station, but are all negatively related to the most recent qualifications at the gunner's station.

Table E-1 presents the correlation of the other biographical measures with success at Annual Qualifications. Since the measures of recency of training cannot be plausibly related to any of the success measures, correlations involving recency of training may be disregarded. As might be expected, the ASVAB composite scores (CO, GT) are positively related to past success at the gunner station. However, for measures of past success at the TC station, the opposite relationship occurs. Similarly, the frequency with which computer games are played is positively correlated with past success at the TC station but negatively correlated with past success at the gunner's station.

With regard to the Computer Panel job sample (cf, Table E-2 and 4-20), the data is consistent in showing that the number correct is positively related to success at the most recent Annual Qualifications at both the TC and gunner's station. Namely, the greater the number correct, the higher the qualification score. These relationships do not hold, however, when one considers the qualification average measures. When completion time is considered, no consistent pattern emerges across either the success measures or the two completion time measures. There is one significant correlation--that relating completion time on the CST task with QAVG:G. In accordance with expectation, the data indicate that shorter completion times on the CST task are associated with higher qualification scores.

Upon examining the correlations with the Computer Tracking job sample (cf, Tables E-2 and 4-20), an interesting finding emerges. The direction of all correlations with measures of success at the gunner's station are all opposite to expectation. Namely, the less the time on target and the greater the tracking error, the higher the qualification score. Only one of the correlations involving these relationships achieved significance--that between EASY:ERROR and MRQ:G. The opposite effect with Computer Tracking for qualification at the gunner's station are understandable when it is remembered that the Computer Tracking job sample is negatively related to experience. Inconsistent with the above finding, two-thirds of the correlations with qualification at the TC station are in the expected direction. However, the predicted relationships occur consistently only for tracking error. For both measures of success associated with qualification at the TC station, tracking error is negatively related to success in all cases. None of these relationships is significant, however.

For the Computer Target Engagement job sample (cf, Tables E-4 and 4-16), consistency in the direction of the effect across success measures is achieved only in the case of the time measures. As expected, with only one exception, the shorter the time spent in the 3X and 10X segments of the CTE job sample, the higher the qualification scores. Also as expected, for measures of success associated with qualification at the gunner's station, small target distance error on the CTE task is associated with higher qualification scores in three out of four cases. The relationships between target distance error and success at the TC station, however, is in general in the opposite direction. Despite these trends, none of the correlations with the CTE job sample measures approaches significance.

Correlations with the TC Decision Making job sample follow a consistent interpretable pattern with the effects being much stronger for decision correctness. In three out of the four success measures, contrary to expectation, decision correctness is negatively related to past success. In those cases involving qualifications associated with the TC station, the relationships are significant. The three negative relationships (QAVG:TC, MRQ:TC, QAVG:G) with decision correctness all represent cases where tank crewmen are in general more experienced at making target engagement decisions. If it assumed that the more experienced crewmen have adopted their own strategy for engagement (see previous discussion in Secti[illegible] 4.3.4), then these negative relationships are understandable. The [illegible] cell where a positive correlation was obtained (MRQ:G) represents a sit[illegible]ion where the tank crewmen have, in general, had less experience.

When relationships with decision correctness are examined, the dat[illegible] consistent in exhibiting a negative relationship in each of the fou[illegible] [illegible]ccess measures. Irrespective of the crewstation at which the tank crewmen qualified, short decision times are associated with higher qualification scores. None of these relationships with decision time are significant, however.

Relationships involving the Hands-on Gun Laying job sample vary depending upon whether qualification averages or the most recent qualification scores are utilized in the data analysis (cf., Tables E-5 and 4-20). When qualification averages are considered, the data is consistent in showing that gun laying error is negatively associated with past success as predicted. But contrary to prediction, time to lay the gun is positively associated with past success. This latter finding is understandable if one assumes the better tank crewmen take slightly longer to lay the gun to achieve a higher degree of accuracy. In spite of the consistent pattern with qualification averages, none of the correlations is significant.

When the most recent qualification scores are examined, a different pattern of relationships emerge with the gun laying measures. For the most recent qualification when in the TC station, there is a slight trend for short gun laying times and less gun laying error to be associated with high qualification scores. However, for those tank crewmen qualifying at the gunner station, these same relationships are in the opposite direction, with the correlation between GL:TIME and MRQ:G reaching significance.

For the Hands-On Tracking job sample, different patterns emerge for qualifications associated with the TC and gunner's station (cf., Tables E-6 and 4-20). When at the TC station, tracking distance is positively related to past success as expected. However, contrary to expectation, the number of tracking hits is negatively correlated with past success at the TC station. When qualifications at the gunner's station are considered, findings with regard to tracking distance are opposite to prediction and opposite to that for qualification at the TC station. Namely, short tracking distances are associated with higher qualification scores at the gunner's station. Finally, for the relationships involving the number of tracking hits with past success at the gunner station, the effects are opposite for the average and the most recent qualification measures.

In exact correspondence with the Hands-On Tracking job sample, the pattern of the results for the Hands-On Target Engagement job sample varies with crewstation. Namely, when qualification was at the TC station, short gunner engagement times tended to be related with higher qualification scores. But contrary to expectation, the number of target engagement hits was negatively related to qualifications when in the TC station. Again contrary to prediction and inconsistent with the findings at the TC station, gunner engagement time was positively related to qualification scores at the gunner station. When the number of engagement hits is considered, there is a positive relationship to success at the gunner's station only for the most recent qualification. In spite of the above trends, the only relationship to achieve significance was that between TOT:HITS and MRQ:TC.

To recapitulate, the bivariate relationship between the job sample measures and the measures of past success are not very strong, are largely nonsignificant, and vary across crew qualification position (TC versus gunner) and qualification measure (QAVG versus MRQ).

4.5.2 Multiple Correlations with Past Success

To ascertain if individual job samples as a whole predict past success, separate multiple correlations with past success were computed for each job sample. Only the primary job sample measures were included in the multiple correlation with QAVG:TC, QAVG:G, MRQ:TC, and MRQ:G. In addition, to provide a baseline for comparison, multiple correlations with the Army experience measures were calculated as well.

In the experience-related multiple correlations, only four of the five experience measures were included. A3:TIME was dropped from the analysis because very few subjects had spent any time in the M60A3 and to provide comparability with the results presented in the next section.

Table 4-21 presents the results of the multiple correlations with the four measures of past success. In examining these correlations, attention should again be focused upon the data for the combined subsamples. It should also be pointed out that the magnitude of the multiple correlations are affected by the number of variables included in the prediction subset. The greater the number of variables, the greater the likelihood of obtaining higher multiple correlations; therefore, the magnitude of the multiple correlation cannot be directly compared across job samples.

Inspection of Table 4-21 reveals that the relationships which emerge with the job samples as a whole are generally stronger than was the case for the bivariate relationships. For qualification at the TC station, the multiple correlations range from .087 to .550 with the highest correlation being between the TCD job sample and the most recent qualification success measure. This latter finding indicates that the TCD job sample alone explains

TABLE 4-21. MULTIPLE CORRELATIONS BETWEEN MEASURES OF ARMY EXPERIENCE AND OF PERFORMANCE ON 7 JOB SAMPLES AND MEASURES OF PAST SUCCESS AT ANNUAL QUALIFICATIONS (FROM BIOGRAPHICAL DATA) FOR TANK COMMANDERS (TC), GUNNERS (G), AND COMBINED (TCG) SUBSAMPLES

Experience/ Job Sample	Variables in Set	Sample	QAVG: TC	MRQ: TC	QAVG: G	MRQ: G
Army Experience	RANK	TC	.349(31)[1]	.240(23)	.344(25)	---(3)
	ARMY:TIME	G	.459(7)	.352(6)	.190(33)	.203(21)
	A1:TIME	TCG	.435(38)	.357(29)	.174(58)	.195(24)
	CP:TIME					
Computer Panel (CP)	ECD:CORR	TC	.312(25)	.341(19)	.591(20)	---(3)
	ECD:TIME	G	---(5)	---(4)	.341(27)	.212(18)
	CST:CORR	TCG	.244(30)	.368(23)	.428(47)	.168(21)
	CST:TIME					
Computer Tracking (CT)	EASY:TOT	TC	.572(23)	.429(17)	.467(19)	---(3)
	EASY:ERROR	G	---(6)	---(5)	.324(25)	---(16)
	MOD:TOT	TCG	.457(29)	.349(22)	.306(44)	.895(19)**
	MOD:ERROR					
	HARD:TOT					
	HARD:ERROR					
Computer Target Engagement (CTE)	PROC:ERROR	TC	.347(29)	.179(22)	.377(23)	---(3)
	3X:ERROR (AVG)	G	.991(7)	---(6)	.506(32)	.310(20)
	3X:TIME	TCG	.396(36)	.215(28)	.306(55)	.316(23)
	10X:ERROR (AVG)					
	10X:TIME					
TC Decision Making (TCD)	D:CORR	TC	.284(32)	.612(24)*	.340(25)	---(3)
	D:TIME	G	.769(7)	.916(6)	.081(33)	.273(21)
		TCG	.328(39)	.550(30)**	.206(58)	.263(24)
Hands-On Gun Laying (HGL)	GL:TIME	TC	.364(31)	.280(24)	.316(25)	---(3)
	GL:ERROR	G	.354(7)	.479(6)	.305(32)	.642(21)**
		TCG	.184(38)	.087(30)	.247(57)	.619(24)**
Hands-On Tracking (HT)	TC:HITS	TC	.275(31)	.300(24)	.348(25)	---(3)
	TC:DIST	G	---(6)	---(6)	.395(33)	.347(21)
	G:HITS	TCG	.247(38)	.231(30)	.156(58)	.324(24)
	G:DIST					
Hands-On Target Engagement (HTE)	TOT:HITS	TC	.144(28)	.355(21)	.140(22)	---(2)
	TC:TIME	G	.918(7)	---(5)	.479(30)	.383(18)
	G:TIME	TCG	.324(35)	.525(26)	.184(52)	.364(20)

* $p \le .05$
** $p \le .01$

[1]Number in parentheses indicates the sample size upon which each multiple correlation is based,

30 percent of the variation in the qualification scores at the TC station for the most recent Annual Qualification. Somewhat disappointing, however, is the fact that the combination of experience measures in general accounts for more variability in the TC success measure than do individual job samples. For QAVG:TC, only one job sample (i.e., CT) accounts for more variability than does Army experience, whereas for MRQ:TC, three job samples (i.e., CT, TCD, and HTR) account for more variability.

For qualification at the gunner station, the multiple corelations ranged from .156 to .895. The highest multiple correlation, that between the computer tracking job sample and MRQ:G, indicates that the CT job sample alone explains 80 percent of the variability in the most recent qualification score at the gunner's station. When compared to the variability explained by the experience measures, six of the seven job samples have higher multiple correlations with past success at the gunner's station. This latter finding holds for both gunner qualification measures.

Despite the fact that the multiple correlations are higher than the bivariate correlations, significance is achieved only for the most recent qualification measures. For the most recent qualification at the TC station, only the TC Decision Making job sample accounts for a significant proportion of the variability. Given that significance is found for a job sample which represents a task normally performed by tank commanders, the construct validity of the TCD job sample is enhanced.

For the most recent qualification at the gunner's station, two job samples explain a significant proportion of the variability--Computer Tracking and Hands-On Gun Laying. The finding of a significant relationship between the HGL job sample and qualification scores at the gunner's station and not the TC station was quite unexpected. This is even more puzzling when it is remembered that the direction of the relationships between the HGL measures and MRQ:G are opposite to expectation. Namely, greater gun laying error and longer gun laying times are associated with higher qualification scores.

4.5.3 Multiple Regression of Job Sample Measures with Past Success

The results involving the bivariate and multiple correlation with the past success measures are somewhat disappointing in that the relationships were in general not very strong. To provide some indication of the criterion related validity of the job samples, forward stepped multiple regression analyses were performed incorporating variables from all job samples.

For each of the four post success measures, four different multiple regression analyses were conducted. For each success measure, two analyses involved listwise deletion of missing cases and two involved pairwise deletion of missing cases. With listwise deletion (the typical procedure), regression analysis is performed only on those cases for which there is a complete set of data. In listwise deletion, only correlations based upon valid cases are utilized in the data analysis. As a consequence, the sample size is greatly reduced and the correlations among the predictor (job sample) measures are not what have been presented thus far in this report. Thus, although listwise deletion regressions do validly represent the restricted set of data they are attempting to explain, they do not represent the relationships heretofore discussed.

Pairwise deletion of missing cases, on the other hand, utilizes the correlations upon which each pair of measures is based in the total data base. This option causes the case to be eliminated from calculations involving that variable only. As a consequence, each correlation is based upon a different sample size. Although the multiple regression analysis with the pairwise deletion option more accurately reflect the intercorrelational structure which exists in the total data base, the results may also be distorted. Since the minimum sample size upon which any correlation is based always involves correlation of the job sample measures with past success, one is using correlations based upon a greater sample size to explain correlations based upon a much smaller subset of subjects.

Since both options have their drawbacks, multiple regression analyses were conducted using both listwise and pairwise deletion. For each of the two deletion methods, residualized and nonresidualized regression analyses were performed. In the residualized regression analyses, the effect of the experience measures was first removed prior to testing the significance of the job sample measures. For the nonresidualized regressions, the explanatory power of individual job sample measures was directly tested without first removing the effect of experience.

Appendix E (Tables E-9 through E-16) contains the results of the 16 regression analyses. These results are summarized in Table 4-22.

Table 4-22 incorporates only the results pertaining to the five variables which add most to the ability to predict each success measure. In that no single job sample had more than five major measures, the number of variables included in this summary table was limited to maximally five so that direct comparisons could be made to the multiple correlation results. In some cases, fewer than five variables are included because all of the variability in the criterion variable had been explained. As can be seen, Table 4-22 indicates the best predictor measures, the sign of the bivariate correlations for the best prediction, and the proportion of variability explained by the linear combination of predictors.

In spite of the relatively low correlations with past success on the level of individual measures and the level of individual job samples, the results of the regression analyses indicate that linear combinations of predictor measures across job samples account for a very high proportion of the variability of the four success measures. For qualifications at the TC station, the proportion of variability explained ranges from .582 to .789 for the average qualification measures and from .722 to .881 for the most recent qualification score. In five of the eight analyses, involving qualification at the TC station, the linear combination of predictors accounts for a significant proportion of variability.

TABLE 4-22. REGRESSION ANALYSES OF ANNUAL QUALIFICATION SCORES ON MAJOR MEASURES OF PERFORMANCE FOR SEVEN JOB SAMPLES

		Results of Regression Analyses															
		QAVG:TC				MRQ:TC				QAVG:G				MRQ:G			
		Residualized		Nonresidualized		Residualized		Nonresidualized		Residualized		Nonresidualized		Residualized		Nonresidualized	
Job Sample	Variable	P	L	P	L	P	L	P	L	P	L	P	L	P	L	P	L
CP	ECD:CORR						+		+	−		−					
	ECD:TIME										−						
	CST:CORR				+	+											
	CST:TIME		−		−					−	−	−	−				
CT	EASY:TOT																
	EASY:ERROR								−					+	+	+	+
	MOD:TOT	−						−									
	MOD:ERROR	−								−		−					
	HARD:TOT																
	HARD:ERROR				−		−							+		+	
CTE	PROC:ERROR		+						+								
	3X:ERROR (AVG)	+		+													−
	10X:ERROR (AVG)				+		−			−	−	−	−				
	3X:TIME										+						
	10X:TIME				−								+				−
TCD	D:CORR	−		−		−		−							+		+
	D:TIME														−	−	−
HGL	GL:TIME	+								+		+		+		+	
	GL:ERROR			−				−									
HT	TC:HITS			−		−		−								+	
	TC:DIST		+						+				−		+		
	G:HITS		−														
	G:DIST		+						+								
HTE	TOT:HITS			−		−		−									
	TC:TIME																
	G:TIME						−				+		+				
R-Square Experience		.258	.374			.160	.226			.031	.097			.076	.367		
R-Square Job Samples		.589	.617	.582**	.789*	.836*	.722	.881*	.866*	.327	.707*	.313*	.674**	.820*	.630	.996**	.968**

When qualification at the gunner station is considered, the proportion of variability explained ranges .313 to .707 for the average qualification measure and from .630 to .996 for the most recent qualification score. The linear combination of predictors accounts for a significant proportion of variability in six of the eight analyses representing qualification at the gunner's station. For a seventh analysis, although nonsignificant results were obtained due to the lack of statistical power, 100 percent of the variability had been explained.

Inspection of Table 4-22 reveals that the job sample measures which are the best predictor are contingent upon the qualification measure (average versus most recent), the regression deletion option (listwise versus pairwise), and the variance to be explained (residualized versus nonresidualized). Given the differences which emerged across the above factors, no attempt is made to describe and explain individual predictor criterion relationships.

One very important fact is worthy of mention, however. No single job sample appears to be the best predictor of past success at Annual Qualifications. If the pattern of results in Table 4-22 is any indication, it is likely that a combination of predictor measures across job samples will be necessary to predict future success at Annual Qualifications in Phase II of the present investigation.

Section 5

PHASE I: SUMMARY AND CONCLUSIONS

The summary and conclusions regarding analysis of the predictor data are centered around the three purposes advanced in Section 4.

5.1 INTERCORRELATION AMONG INDIVIDUAL MEASURES OF JOB SAMPLE PERFORMANCE

As was seen in Table 4-18 (cf, Section 4.4.5.1), the bivariate correlations are in general low, with approximately one-half of the correlations being 0.150 or below. Correspondingly, the percentage of bivariate correlations that are significant are also low. In only a few cases do the number of significant correlations exceed chance expectations.

The low degree of intercorrelation among the job sample measures is ideal from a measurement perspective. With a low intercorrelation among the job sample measures, there is a greater probability of finding measures which will add to the ability to predict the criteria variables in Phase II of the present investigation.

5.2 ROLE OF EXPERIENCE IN JOB SAMPLE PERFORMANCE

Section 4.3.8.3 presented a summary of the evidence regarding the role of experience as a factor in job sample performance. As was indicated in Table 4-16, presumed differences in the experience level of tank commanders and gunners resulted in significant differences in job sample performance in only one case--Hands-On Gun Laying. Furthermore, when differences in experience within tank crewmen subsamples are considered, there is statistical evidence for experience being a factor only for the Computer Tracking job sample. In that job sample, quite opposite to expectation, experience is in general negatively related to performance for both tank commanders and gunners; that is, higher experience levels are associated with less time on target and greater tracking error. In the other job samples, measures of experience are not consistently related to performance differences within the tank commander and gunner subsamples.

The fact that there is a lack of statistical evidence for experience as a factor in job sample performance at least makes tenable the assumption that one is measuring aptitudes. However, as discussed in detail in Section 4.1, one cannot definitely affirm that differences in aptitudes are responsible for differences in job sample performance.

5.3 VALIDITY OF THE JOB SAMPLES AND JOB SAMPLE APPROACH IN ABSENCE OF POSTPREDICTION CRITERIA DATA

Two types of validity assessment were involved in the analysis of the predictor data-construct validity and criterion-related validity. To assess construct validity, logical deductions were first made about the direction of relationships between job samples and job sample measures. These were then compared to the obtained correlations to determine if the presumed relationships occurred. To evaluate the criterion related validity of the job samples, job sample measures were correlated with measures of past success at Annual Qualifications.

If a single tank gunnery aptitude underlies job sample performance, then there should be a high intercorrelation among job sample measures, both between and within job samples. In that there were relatively high intercorrelations among the performance measures within job samples but not between job samples, the evidence indicates that the job samples tap different behaviors and possibly different aptitudes.

As summarized in Table 4-18 (cf, Section 4.5.1), the intercorrelation between measures of different job samples were low and with only a few exceptions were nonsignificant. In that there are major task differences between job samples, the low intercorrelations between measures of different job samples were not unexpected.

However, despite the lack of significance, a predictable pattern in the direction of relationships emerged among the job sample measures which cannot be attributed solely to random chance variation. That is, based upon knowledge of tank gunnery and human performance, predictions about the direction of relationships were advanced. In many cases the presumed direction of relationship was confirmed by the data. Table 4-19 (cf, Section 4.5.2) summarized those job sample comparisons in which the predicted direction of relationships emerged.

To recapitulate, there was generally good evidence for the construct validity of the job samples. The job samples were designed to tap different behaviors and that they do was confirmed by the low intercorrelation among measures of different job samples. Despite the low and nonsignificant correlations, the predicted pattern of relationships emerged in the majority of cases. For two job samples in which relationships were definitely expected (i.e., computer tracking and hands-on tracking), the predicted pattern emerged and significance was obtained, but for gunners only. The major piece of disconfirming evidence came from the fact that there was no evidence for a relationship between the two target engagement tasks.

In absence of any postpredictor criterion data, there was also good evidence for the criterion-related validity of the job samples. The results of the regression analyses indicated that linear combinations of predictor measures across job samples account for a very high proportion of the variability in past success at Annual Qualifications. Furthermore, the results also indicate that no single job sample appears to be the best predictor of past success. If the results of the multiple regression analyses are any indication, it is likely that a combination of predictor measures across job samples will be necessary to predict postpredictor success at Annual Qualification in Phase II of the present investigation.

REFERENCES

Black, B. A., & Kraemer, R. E. (1981). XM1 gunnery training and aptitude requirements analyses (Research Proudct 81-5). Alexandria, VA: U.S. Army Research Institute.

Cream, B. W., Eggemeir, F. T., & Klein, G. A. (1978). A strategy for the development of training devices. Human Factors, 20, 145-158.

Eaton, N. K. (1978). Predicting tank gunnery performance (USARI Research Memorandum 78-6). Alexandria, VA: U.S. Army Research Institute.

Eaton, N. K., Bessemer, D. W., & Kristiansen, D. M. (1979). Tank crew position assignment (Technical Report 391). Alexandria, VA: U.S. Army Research Institute.

Eaton, N. K., Johnson, J., & Black, B. A. (1980). Job samples as tank gunnery performance predictors (Technical Report 473). Alexandria, VA: U.S. Army Research Institute.

Guion, R. M. (1979a). Principles of work sample testing: I. A nonempirical taxonomy of test uses (Technical Report 79-A8). Alexandria, VA: U.S. Army Research Institute.

Guion, R. M. (1979b). Principles of work sample testing: II. Evaluation of personnel testing programs (Technical Report 79-A9). Alexandria, VA: U.S. Army Research Institute.

Guion, R. M. (1979c). Principles of work sample testing: III. Construction and evaluation of work sample tests (Technical Report 79-A10). Alexandria, VA: U.S. Army Research Institute.

Guion, R. M. (1979d). Principles of work sample testing: IV. Generalizability (Technical Report 79-A11). Alexandria, VA: U.S. Army Research Institute.

Maier, N. R. F. (1973). Psychology in industrial organizations. Boston: Houghton Mifflin Company.