Performance in Four Army Jobs by Men at Different Aptitude (AFQT) Levels: 4. Relationships Between Performance Criteria

Robert Vineberg and Elaine N. Taylor

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

Approved for public release; distribution unlimited.

August 1972

Prepared for
Office of the Chief of Research and Development
Department of the Army
Washington, D.C. 20310
Destroy this report when it is no longer needed.
Do not return it to the originator.
A study was made of approximately 1800 men with experience ranging to 20 years in five different Army MOSs to provide information about the performance and characteristics of effective and ineffective marginal personnel in the Army. The study included a group of men with Armed Forces Qualification Test scores (AFQT) in the marginal range and a comparison group of men in the same jobs, but in the upper range of AFQT scores. Performance was measured by intensive job sample tests, job knowledge tests, and supervisor ratings. Biographical questionnaires, a battery of published and experimental tests, and Army records provided information about background, personal characteristics, and military experiences. This report, the fourth in a series presenting the extensive data and analyses, examines the determinants of job behavior and describes the relationships among the three performance criteria used in the study: job sample tests, job knowledge tests, and supervisor ratings.
<table>
<thead>
<tr>
<th>KEY WORDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armed Forces Qualification Test scores</td>
</tr>
<tr>
<td>Individual differences</td>
</tr>
<tr>
<td>Individual training</td>
</tr>
<tr>
<td>Job analysis</td>
</tr>
<tr>
<td>Job description</td>
</tr>
<tr>
<td>Job knowledge tests</td>
</tr>
<tr>
<td>Job sample tests</td>
</tr>
<tr>
<td>Marginal personnel</td>
</tr>
<tr>
<td>MOS performance</td>
</tr>
<tr>
<td>Performance criteria relationships</td>
</tr>
<tr>
<td>Project 100,000</td>
</tr>
<tr>
<td>Supervisor ratings</td>
</tr>
<tr>
<td>Test administration</td>
</tr>
<tr>
<td>Test development</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROLE</td>
<td>WT</td>
<td>ROLE</td>
</tr>
</tbody>
</table>
Performance in Four Army Jobs by Men at Different Aptitude (AFQT) Levels:
4. Relationships Between Performance Criteria

Robert Vineberg and Elaine N. Taylor

HumRRO Division No. 3
Presidio of Monterey, California
HUMAN RESOURCES RESEARCH ORGANIZATION

Approved for public release; distribution unlimited.

Prepared for
Office of the Chief of Research and Development
Department of the Army
Washington, D. C. 20310
The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

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.
FOREWORD

HumRRO Work Unit UTILITY was initiated in January 1967 as part of Project 100,000 to provide information about the Army performance and characteristics of marginal personnel, men whose scores on the Armed Forces Qualification Tests were in the low levels. Its objectives were (a) to find out how men in Mental Category IV compared with men in other mental categories in the performance of selected Army jobs, and (b) to identify different factors associated with satisfactory performance in different mental category groupings—specifically, to explore the relationships among a man’s background, personal characteristics, Army experiences, and job performance.

The UTILITY research has been conducted by HumRRO Division No. 3 at the Presidio of Monterey, California, with Dr. Howard McFann as Director. Dr. Robert Vineberg was the Work Unit Leader. Members of the research team at various times have included Dr. Elaine N. Taylor, Dr. John S. Caylor, Miss Annette K. Mahikoa, Dr. S. James Goffard, Dr. Thomas G. Sticht, Dr. Joseph S. Ward, Dr. Herbert G. Gerjuoy, Mr. Donald F. Polden, and Mr. Leon E. Guyton.

Military support for the study was provided by the U.S. Army Training Center Human Research Unit. Military Chief of the Unit at the beginning of the research was LTC David S. Marshall; during the period of data collection and early analysis, LTC Robert J. Emswiler was Unit Chief; the present Chief is COL Ullrich Hermann.

Enlisted men assigned to the project during the data analysis and report preparation phases included SGT Gerald G. Lynch, SP5 William Yanda, and SP4 Gregory Herr.

The extensive findings from this research are being reported in a series of publications. This, the fourth report, examines the determinants of job behavior and describes the relationships among the three performance criteria used in the study: job sample tests, job knowledge tests, and supervisor ratings. The first report in the series described the rationale, research design, and general chronology of research events. The second report described the data collection instruments used in the study and their development and administration. The third report presented the bulk of the major study findings, including comparisons of the performance of men in different mental categories with different amounts of job experience; comparisons of the performance of special subgroups (Blacks and Whites, Inductees and Enlistees, and Men with Formal and On-the-Job Training); an analysis and definition of acceptable performance; a procedure for using job knowledge tests to screen ineffective performers; and an analysis of the relationship between personal characteristics and job performance.

HumRRO research for the Department of the Army is conducted under Contract DAHC 19-70-C-0012. Training, Motivation, Leadership Research is conducted under Army Project 2Q062107A712.

Meredith P. Crawford
President
Human Resources Research Organization
INTRODUCTION

Work Unit UTILITY was concerned with a comparison of the job proficiency of men in Mental Category IV and other mental categories in four different Army jobs: Armor Crewman, General Vehicle Repairman, Unit and Organizational Supply Specialist, and Cook.¹

PROBLEM

A fundamental problem in the assessment of job proficiency is the selection of appropriate and satisfactory criterion instruments. The type of measure to be used depends upon a variety of considerations, including the purpose and situation in which assessment is being undertaken, the feasibility of employing a specific type of measure in a particular operational or research setting, and the costs in time, personnel, and money that are required. Because practical constraints inevitably can be expected to promote the use of criterion instruments that are simplest to administer, it becomes essential to examine both the qualities needed in criterion instruments and the relationships between different types of criterion instruments to determine where substitutions of one for another can reasonably be considered. This report is designed to provide the reader with an analysis and a general review of the factors that govern the proper selection of criterion measures for different situations and purposes.

In UTILITY three different kinds of criterion instruments were used: job sample tests, job knowledge tests, and supervisor ratings. In general, these three classes of criterion measures differ conceptually in their approach to the assessment of job proficiency, and, as a consequence, can focus on quite different elements of job behavior. Job sample tests are based upon samples of the performances required in a job. Job knowledge tests are based upon an analysis, extraction, and sampling of the knowledge required in job performance. Supervisor ratings are based upon verbal descriptions of job behavior or determinants of job behavior. The data collected in UTILITY provide a means for an empirical examination of the relationship among these three different varieties of criterion performance in four different jobs.

APPROACH

In the selection of a criterion instrument, a general understanding of the primary determinants of job behavior is required. These are:

Job Knowledge—Information about material used on the job and about acts, procedures, and principles needed to mediate job performance.

Job Skill—Performances not mediated completely by knowledge or conscious verbal processes (self-instruction) and not readily acquired solely through verbal instruction. Skills include perceptual, motor, cognitive, and social skills.

¹Information about the Medical Specialist, a fifth MOS included in the study, is not contained in this report. Job specialization in this MOS was found to be considerable. Sampling procedures and analysis of these data differed from those followed for the other four jobs, and for this reason information on this MOS has not been included here.
Motivation and Personality Characteristics—Attitudes, values, and general styles of behavior that condition the manner in which knowledge and skill are employed in job performance.

Selection of an appropriate criterion instrument depends upon the relative contribution of each of these factors to performance in a specific job situation. Selection of a criterion instrument also depends upon the purpose for which measurement is being undertaken. The major purposes of job proficiency measurement can be summarized as follows:

To provide estimates of the effectiveness of an individual’s job performance.
To diagnose deficiencies during training or on the job.
To evaluate the effectiveness of a training program.
To provide estimates of group performance in research studies.

The rationale that underlay the selection of the three types of criterion instruments (job sample, job knowledge, and supervisor ratings) used in UTILITY is presented in the report.

As observed earlier, practical constraints often dictate that relatively simple and inexpensive measures be used to assess job proficiency, sometimes at the sacrifice of more valid and appropriate information. The report explores the interchangeability of job knowledge and job sample tests, and the considerations that govern the suitability of such substitution. An analysis was made of the skill and knowledge requirements of the job sample tests that are considered to exemplify the demands of the jobs themselves. The analysis suggests that well-constructed knowledge tests—that is, tests that contain only information required in actual job performance—could be substituted for job sample tests, at least in the four jobs studied.

The reliabilities and interrelationships of job sample tests, job knowledge tests, and supervisor ratings were then determined. These data support the analyses and conclusions already reached.

The final step in the review was a study of the relationships between job sample tests and job knowledge tests for selected subsamples: low vs. high reading ability groups, Mental Category IV vs. Non-IV groups, low vs. high verbal aptitude groups, low vs. high educational level groups, and black vs. white groups.

RESULTS AND CONCLUSIONS

(1) Correlations between job sample test scores and job knowledge test scores in the four jobs ranged from .58 to .72. Job knowledge test reliability coefficients ranged from .76 to .92. Estimates of the lower limit of job sample test reliabilities ranged from .35 to .55.

(2) Correlations between job sample test scores and supervisor ratings ranged from .20 to .28. Correlations between job knowledge test scores and supervisor ratings ranged from .23 to .35.

(3) Analysis of the job sample tests used to measure performance for the four jobs in this study—Armor Crewman, Repairman, Unit and Organizational Supply Specialist, and Cook—showed that the skill requirements in these jobs are minimal. Therefore, job knowledge tests could appropriately be substituted for more direct measures of performance such as job sample tests in assessing proficiency in these jobs.

(4) The selection of appropriate criterion measures for assessing job proficiency requires that the skill and knowledge components of a job be identified. Where job
performance is mediated almost solely by knowledge per se, job knowledge tests should be used to assess the information a man has about his job. Where job performance relies almost solely on skill, job sample tests or some other variety of performance measure is essential. Where job performance depends upon a mix of knowledge and skill, both knowledge tests and performance tests are indicated. An instrument that has been designed to assess job motivation and other personal qualities, such as rating scales, is always essential if a complete estimate of job performance is desired.
CONTENTS

Introduction ................................................. 3
  Background ............................................. 3
  Objectives of the Study ................................. 4

Components of Job Performance ......................... 5
  Job Knowledge ........................................... 5
  Job Skill ................................................ 6
    Perceptual Skills ..................................... 6
    Motor Skills .......................................... 7
    Cognitive Skills ...................................... 7
    Social Skills ......................................... 8

Motivation and Personality Characteristics ............ 9

Types of Proficiency Measures and Purposes of Measurement ......................... 10
  Measures to Assess Job Proficiency .................... 10
  Purposes of Measurement ................................ 11
  Description of the UTILITY Tests ..................... 12

An Analysis of the Interchangeability of Proficiency Measures and Evidence From the UTILITY Data ......................... 16

Interrelationships of Job Sample Tests, Job Knowledge Tests, and Supervisor Ratings ......................... 18

Relationships Between Job Sample Tests and Job Knowledge Tests in Selected Subsamples ......................... 21

Conclusion .................................................. 24

Literature Cited ........................................... 29

Tables
  1 Job Sample Test ........................................... 13
  2 Frequency of Steps in Job Sample Tests Categorized According to Knowledge and Skill Requirements ......................... 18
  3 Correlations of Job Sample, Job Knowledge, and Supervisor Ratings for the Entire Sample ......................... 19
  4 Job Sample Test Reliabilities (Lower Limits) ......................... 19
  5 Job Knowledge Test Reliabilities (Correlations Between Odd and Even Numbered Items) ......................... 20
  6 Correlations of Job Sample and Job Knowledge for Low Reading and High Reading Ability Subgroups ......................... 22
  7 Correlations of Job Sample and Job Knowledge for Category IV and Non-Category IV Subgroups ......................... 22
  8 Correlations of Job Sample and Job Knowledge for Low Verbal and High Verbal Aptitude Subgroups ......................... 23
### Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Correlations of Job Sample and Job Knowledge for Low Education and High Education Level Subgroups</td>
<td>23</td>
</tr>
<tr>
<td>10</td>
<td>Correlations of Job Sample and Job Knowledge for Black and White Subgroups</td>
<td>24</td>
</tr>
</tbody>
</table>

### Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Steps in Fan Vertical Shaft Oil Seal Replacement Subtest—Repairman Job Sample Test</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Illustrative Items From the Supply Specialist Job Knowledge Test</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>Illustrative Items From the Supervisor Rating Scales (Enlisted Efficiency Report)</td>
<td>16</td>
</tr>
</tbody>
</table>
Performance in Four Army Jobs by Men at Different Aptitude (AFQT) Levels:
4. Relationships Between Performance Criteria
INTRODUCTION

BACKGROUND

The most fundamental and difficult problem in any attempt to measure job proficiency is the selection of appropriate and satisfactory criteria. The choice of a particular measure or criterion of proficiency depends upon a variety of considerations, including the purpose or purposes for which assessment is being undertaken, its feasibility for use in a particular operational or research setting, and its cost. An important factor determining the suitability of a criterion is the extent to which it measures aspects or determinants of behavior that differentiate among performers in a particular job with regard to factors that are critical determinants of performance in that job.

This report examines the suitability of different criteria in relation to the specific behavioral demands of a job. In particular, it focuses on the usefulness and relationships among three types of criterion instruments: job sample tests, job knowledge tests, and supervisor ratings. Each of these instruments has been used to assess men working in the four Army jobs—Armor Crewman, General Vehicle Repairman, Unit and Organizational Supply Specialist, and Cook—that were included in Work Unit UTILITY, a comparative study of the performance of men in different mental ability groups.

The primary objectives of the overall UTILITY research program were to find out how men in Category IV\(^1\) and other mental categories compare in the performance of selected Army jobs, and to identify different factors associated with satisfactory performance in different mental category groupings. The purpose and design of Work Unit UTILITY have been described in an earlier report in this series (1).

Other reports have described the data collection instruments used in the study, including the criterion measures (2), and job performance as it varies both in accordance with the main study variables—AFQT level and amount of job experience—and in different ethnic groups (3).

In general, the three different kinds of criterion instruments used in the UTILITY research—job sample tests, job knowledge tests, and supervisor ratings—differ conceptually in their approach to the assessment of job proficiency. As a consequence, they can focus on quite different elements of job behavior:

1. Job sample tests are based upon samples of performance required in a job. Where performance is mediated almost entirely by knowledge, a valid job sample test will necessarily measure knowledge. Where performance is mediated almost entirely by skill, a valid job sample test measures largely skill. Where job performance is mediated by both knowledge and skill, valid job sample tests measure both knowledge and skill.

2. Job knowledge tests are based upon an analysis, extraction, and sampling of the knowledge required in job performance and are intended solely to measure knowledge. They cannot be used to measure skill.

3. Supervisor ratings are based upon verbal descriptions of job behavior or determinants of job behavior. Although supervisor ratings can be applied to the

---

\(^1\)This includes men in Mental Categories IV and V (i.e., men who received a percentile score of 0-30 on the Armed Forces Qualification Test, (AFQT). Men in these categories are considered marginal and their acceptance into the military depends upon the minimum standard prevailing at the time.
measurement of many different aspects of job behavior, they generally provide more
global estimates of performance, such as attitude toward the job, and general personal
adaptability and reliability.

In the UTILITY research, job sample tests were viewed as providing the closest
approximation to the ultimate criterion of actual job performance. Job knowledge tests
were used because they provide a very efficient means of sampling from the entire range
of knowledge required in job performance. Supervisor ratings were used because they
provided the only feasible means for assessing motivation and other personal
characteristics that a job incumbent had demonstrated in his job. Also, these three
different types of criterion measures were included in the research to permit a study of
the interrelationships among the measures.

It is generally recognized that different criteria vary in their suitability for measuring
different components of performance, that criteria vary in their degree of relevance for
assessing proficiency in given jobs, and that one variety of criterion performance is not
necessarily closely related to another. Such understanding is, however, most often
honored in the breach. For example, in the Army even in a job—image interpretation—
requiring extensive perceptual skill, a man's proficiency is assessed on the basis of a
composite of his scores on verbal or paper-and-pencil tests of fact and principle and a
supervisor's ratings of his personal characteristics.¹

Reasons for using tests of information or knowledge or summary ratings of job and
general effectiveness in situations where tests of performance are far more appropriate
are, of course, not difficult to find. Job sample tests are not convenient; they are
expensive in money, in time, and in the professional skill required in test development
and supervision of the testing program. The use of job sample tests in operational
assessment programs, therefore, is often not practicable.

Because practical constraints inevitably can be expected to promote the use of the
criterion instruments that are simplest to administer, it becomes essential to examine
both the qualities needed in criterion instruments and the relationship among different
types of criterion instruments in order to determine where it is reasonable to consider
substitutions of one for another.

OBJECTIVES OF THE STUDY

The data collected in Work Unit UTILITY provide a means for examining the
relationship among different varieties of criterion performance in four fairly different
types of jobs. Armor Crewman job duties, dictated in a large degree by equipment
demands, are highly procedural; the tasks may be considered machine-ascendant, in that
job behaviors principally involve sequences of interaction between man and machine, in
which a man's activities are largely determined by characteristics of the equipment.
General Vehicle Repairman is a mechanical maintenance job requiring diagnostic and
interpretive behaviors. Supply Specialist is a clerical job calling primarily for the analysis
and recording of information. The Cook's job typically requires reading and following
procedures specified in a cookbook; recognition of standards and precision in meeting
them are primary characteristics of the Cook's job.

The analysis of criterion test selection factors includes the following:
(1) An analysis of the primary determinants of job behavior.

¹Department of the Army, "Proficiency Pay for Army Enlisted Personnel," Chapter 6, Enlisted
A classification of proficiency measures, specification of the major purposes of proficiency measurement, and discussion of the rationale underlying the selection of the three instruments used in UTILITY.

A discussion of the interchangeability of job sample and job knowledge tests and an analysis of the skill and knowledge requirements in the four jobs studied in UTILITY.

A presentation of the reliabilities and interrelationships of the three criterion instruments.

A presentation of the relationships between job sample tests and job knowledge tests for selected subsamples of the study.

COMPONENTS OF JOB PERFORMANCE

A first step in relating the suitability of different types of criterion measures to the characteristics of jobs is the identification of relevant components of performance that can be used to classify jobs. Several investigators have undertaken the classification of tasks, training objectives (tasks to be learned), and jobs (4, 5, 6, 7, 8, 9, 10, 11). Some of these efforts, such as that of Fine (5), are quite comprehensive and detailed. For present purposes, a relatively simple set of job components or determinants has been adopted:

(1) **Job Knowledge.** Information about material used on the job and about acts, procedures, and principles needed to mediate job performance.

(2) **Job Skill.** A rather loose grouping of different varieties of behavior not mediated completely by knowledge or conscious verbal processes (self-instruction) and not readily acquired solely through verbal instruction. Thus, job skill is usually thought of as being able to do something rather than knowing what to do. Included are perceptual skills, motor skills, cognitive skills, and social skills.

(3) **Motivation and Personality Characteristics.** Attitudes, values, and general styles of behavior that condition the manner in which knowledge and skill are employed in job performance.¹

JOB KNOWLEDGE

Information about a job generally can be considered the single most important mediator of job behavior. A knowledge of specific acts and procedures guides the worker through the performance of most work activities. Knowledge in the form of facts, principles, and relationships (sometimes in conjunction with the application of cognitive skills) enables an incumbent to respond to contingency situations and to accommodate to situations where procedures must be modified, or where novel forms of response are

¹Omitted from the set of performance components are such determinants of behavior as aptitude and physical ability. Aptitudes are general abilities that condition the acquisition of knowledge and the development of skills. This analysis focuses on knowledge and skill requirements in performance rather than upon factors affecting the development of these capabilities in the performer. Physical abilities are characteristics of physical structure and function that permit job acts to be carried out. While clearly an important determinant of performance, physical ability has been omitted from the set of behavior determinants since most men generally possess the physical abilities required to perform the majority of jobs. Where special needs exist (e.g., tea taster, stevedore, athlete) such factors would obviously have to be represented among the criteria for assessing proficiency.
required. Job knowledge may exist as information that the incumbent brings to the job, or it may be provided in the form of job instructions or job aids.

JOB SKILL

Job skill refers to a heterogeneous grouping of abilities. These abilities are inferred on the basis of behaviors that are not completely mediated by knowledge or verbalization of a to-be-performed act. The acquisition of skill—being able to do something rather than knowing what to do—unlike the acquisition of knowledge, requires practice or rehearsal. Indeed, the need for practice in its development can be considered the major definitional characteristic of a skill. This characteristic is particularly important, because it provides a basis for deciding whether or not skill is being manifested in any given behavior, a decision that is critical in the selection of appropriate criterion instrumentation. Thus, just as the description or specification of skilled behavior is insufficient to mediate performance of that behavior, so expression of knowledge of skilled behavior by an incumbent (e.g., job knowledge test) is insufficient reason for concluding that the skilled behavior can be demonstrated.

A related and common, although probably not essential, characteristic of skill is that skilled behavior typically is not completely specifiable by verbal means (although the ends of such behavior are). For example, a person cannot be told by verbal instruction how to recognize subtle shadings in pitch in auditory tones or in a spoken language or how to swing a baseball bat. These behaviors can be demonstrated, but cannot be completely reduced to verbal description, primarily because the mechanisms underlying them are internal and nonverbal (auditory processes in the case of discriminating tones, kinesthetic processes in swinging a bat), and language generally does not provide commonly agreed upon verbal labels for describing such processes.

Four main classes of job skill can be identified: perceptual skills, motor skills, cognitive skills, and social skills. While for some purposes it may be useful to distinguish additional varieties of skill, such as artistic skills, these four classes seem to cover the critical varieties of skill in most jobs.

Perceptual Skills

Perceptual skills involve the ability to discriminate, interpret, and organize sensory information (e.g., recognizing and interpreting camouflaged targets in aerial photographs, or receiving International Morse Code). The development of perceptual skills such as those involved in photo interpretation requires practice during which a person learns to discriminate varieties and combinations of subtle visual cues. Also, with practice, the identification process undergoes a shortcutting or dropping out of verbal self-instruction or cue-naming that initially aids in mediating performance.

This process can be more clearly observed in the development of skill in receiving code. Here the learner first verbalizes the name of each letter to himself as he hears its signal. He then records the letter. With practice, the naming step drops out and the auditory signal becomes a cue for the direct writing of letters.

The learning of Morse Code also demonstrates another common feature of skill acquisition. With practice, the learner organizes the incoming signals into larger and larger groupings or hierarchies. Initially he attends to signals which denote letters; subsequently, he attends to and “hears” a longer sequence of signals that signify words and ultimately phrases.
Motor Skills

Motor skills (e.g., tracking and leading a moving target with a rifle, or shifting from neutral in an automobile in which the simultaneous letting out of the clutch and depression of the accelerator are coordinated) involve the ability to make appropriate physical responses. During the development of motor skills, a variety of typical processes may occur. The coordination and timing of several simultaneous or sequential movements may be required, as in shifting gears in an automobile. The connection and smoothing of responses into a rhythmic sequence may be needed, as in dancing. The combination and linking of a series of initially discrete responses and adjustments may be necessary, as in combining into a single act the elements involved in returning a tennis ball. As is the case for many perceptual skills, the development of motor skill often requires learning to make rapid and automatic responses (without self-instruction on the part of the performer) or precise or delicate manual movements.1

Cognitive Skills

Cognitive skills involve those covert mental processes associated with the development of concepts or rules, understanding the implications of relationships, and applying rules and principles in making decisions and solving problems. In the present analysis, behaviors have been designated as skilled when knowledge or a description of the behavior in question is not sufficient to ensure its performance and where practice is necessary to produce certain changes and refinements in initial attempts to perform.

When cognitive behavior is examined, one finds the acts or processes themselves almost impossible to describe. Such behavior is only poorly understood. The lack of an external referent to its functioning and the absence of language for describing its underlying processes make it particularly resistant to description. It is difficult to describe how hypotheses are generated other than by way of illustration. A person can be told how to go about identifying his opponent's likely moves in a chess game or how to select electronic components for testing in a troubleshooting task, but he cannot be told how to develop hypotheses in the abstract, mainly because it is impossible to specify the dimensions of anything but a concrete situation.

Because cognitive processes cannot be described adequately or reduced to statements of how cognitive acts are to be carried out, the question of whether the performance of acts of cognition requires something more than knowledge of those acts is largely unanswerable. We are unable to construct a knowledge test consisting solely of questions about how cognitive acts are accomplished, therefore, possession of knowledge of the form of a cognitive act cannot be separated from the performance of the act itself.

Because a knowledge test cannot be built, it is not possible to establish the relationship between cognitive knowledge and the actual performance of cognitive acts. It is appropriate, therefore, to treat cognitive behavior as a skill not because cognitive performance necessarily requires something more than knowledge of how cognitive acts are accomplished, but because knowledge cannot be demonstrated apart from performance. Because knowledge of cognitive acts cannot be independently described and measured, the assessment of cognitive skill will necessarily involve the direct performance of cognitive acts.

The role of practice or training in the acquisition of cognitive skill is also poorly understood. Whether or not practice is always essential in the development of such skill is not clear. Nevertheless, cognitive skills do improve with practice, and typically the only

1As indicated earlier, skill development presupposes certain aptitudes and physical abilities; practice in and of itself is not necessarily sufficient. For example, women seem more likely to possess the manual dexterity required for the soldering of miniaturized equipment; some adults appear unable to recognize subtle inflections in a foreign language regardless of how much training they receive.
means available for effecting this development is to expose persons to situations where
cognition is required. The means by which a person learns to process information, see
relationships, and make inferences and deductions in order to make decisions and solve
problems are not known. Thus, training to develop cognitive skill is inevitably reduced to
a series of demonstrations whereby a learner is exposed to examples of problems and
given practice in dealing with them. In this sense, the development of cognitive skill is
not dissimilar from the development of other varieties of skill.

Social Skills

Social skills involve the ability to work successfully with individuals and groups of
people. In the present analysis, social skills refer to any variety of interpersonal skill and
behavior that is necessary to accomplish a specific job objective. While these skills may
come into play in any work situation involving two or more persons, the skills associated
with effective supervisory performance and effective group performance are of particular
importance.

Social skills have been grouped with the other varieties of skill because precepts or
general principles (knowledge) for effective social interaction do not necessarily lead to a
recommended action. A complete set of principles for effective social interaction is not
available, and existing principles are nonspecific. Rules for identifying situations in which
they apply and the behaviors they call for are not clearly defined. Thus, in responding to
an exhortation such as “be firm yet sensitive to the feelings of others,” it is not certain
what weight is to be given to these somewhat contradictory aspects of behavior. The
specific behavior called for in accommodating to both perceived interpersonal needs
(affective requirements) and job demands (production requirements) is usually not clear.

Furthermore, the ability to apply or act in accordance with a principle often
depends upon the attitude and personality of the individual in question. For example, an
autocratic person may find it impossible to act in a desired laissez-faire or democratic
manner.

Thus, it can be seen that, in general, social skill requires problem solving ability. The
nature of specific personal and situational characteristics must be identified and related to
general principles before action can be taken. Such ability, like the development of
cognitive skill in general, is only acquired through practice in specific situations. Acquisi-
tion of interpersonal skill probably occurs in very small increments, because problem
situations do not completely repeat themselves and feedback on the efficacy of a solution
is, at best, imperfect. More practice is probably required for the development of
interpersonal skill than for any of the preceding varieties of skill.

The acquisition of social skills encompasses the learning of many different things.
People learn to give and take criticism and direction. They learn to role play or see things
from another person’s point of view and to overcome strong personal dispositions. They
learn certain general requirements in dealing with people; individuals must be shown
tolerance and respect, and one must be sensitive to the feelings and needs of coworkers,
superiors, and subordinates.

Some people have strong needs to achieve, to be liked, to be dependent upon or
dominate others; some are aggressive, while others are shy and retiring. Social skill is
dependent upon one’s ability to recognize and accommodate to such needs in others.
People learn to exert a degree of balance or focus in adjusting their behavior in response
to these needs in accordance with specific requirements of the work situation. In a job,
decisions must be made regarding the extent to which a coworker’s personal needs should
influence behavior and the extent to which the need to accomplish specific tasks or ends
should influence behavior.

Actions will vary according to the structure of the social situation. When one is
interacting with a single individual, greater emphasis may be given to accommodating to
his particular style of behavior. When a group of people is involved, other considerations may assume greater importance. One must develop the ability to see the informal structure of a group and to perceive social reality in an undistorted fashion. A new person entering a group, whether he is a supervisor or another worker, needs to recognize the natural leader of the group, existing friendship patterns, and group norms. Problems arise when such features are not recognized and one attempts to respond solely to the formal characteristics of a group.

Most important, a person must develop an awareness of how his own behavior and manner of responding affects others, whether or not he and his feelings are understood as he intends them. Thus, a supervisor may intend to create an atmosphere of understanding and permissiveness when, in fact, he has accomplished precisely the opposite. In some cases the absence of specific guidance may prove far more threatening than a clearly structured and admittedly autocratic work situation. With some workers, attempted permissiveness may be interpreted as a sign of weakness and a lack of ability on the part of the supervisor. Such misunderstanding is more likely to arise if the supervisor does not, in reality, want to play such a role and does not do it well.

A person learns what to observe or look for in interpersonal behavior and how to interpret it. Because the spoken word often cannot be taken at face value, he learns to interpret more subtle actions, intonations, and modes of expression. He also learns to sense or see through biased or distorted information and to anticipate signs of interpersonal conflict revealed in such symptoms as tardiness, avoidance behavior, and critical comments.

As indicated, practice as opposed to the acquisition of information, is crucial for the development of social skill. Learning how others respond and how one should act with others in order to produce certain desired behavior in them can be accomplished only through rather extended exposure to interpersonal situations and problems. Like other skills, social skill can be demonstrated and assessed only through performance in interpersonal situations.

MOTIVATION AND PERSONALITY CHARACTERISTICS

The terms motivation and personality characteristics are used here to refer to a class of more general attitudes, traits, and styles of behavior that condition the manner in which a job will be performed. Included in this class are such characteristics as punctuality, dependability, ability to tolerate frustration, industriousness, and need to achieve.

Successful job performance requires not only that an incumbent possess certain knowledge and skill, but that he possess the motivation, the desire, and the willingness to perform. Different levels of job motivation will result in different levels of job performance. In some jobs, heightened motivation may compensate in part for deficits in knowledge and skill. Also, certain habitual or characteristic ways in which a person generally responds to persons and things may affect his job performance (personality). For example, a person who typically responds to other persons in an aggressive manner is not likely to perform well in a supervisory position.

Although job motivation and personality contribute to job performance, generally it is not practicable to attempt to assess such factors as part of the measurement of job knowledge or job skill. While job motivation and certain personality traits may affect performance during knowledge and skill testing, such testing provides neither an independent nor a reliable estimate of these factors. A man may lack motivation to

Where testing is undertaken for diagnostic purposes, independent assessment of each component of performance is desirable.
work at his job and may perform poorly in it, but he may be highly motivated to perform well in a test situation, and, consequently, do well on the test. Adequate assessment of motivation and personality traits is more effectively accomplished through peer or supervisor ratings based on continuous observation over an extended period of time.1

TYPES OF PROFICIENCY MEASURES AND PURPOSES OF MEASUREMENT

MEASURES TO ASSESS JOB PROFICIENCY

The criterion instruments used in the UTILITY research, job sample tests, job knowledge tests, and supervisor ratings, may be identified among a large variety of measures that can be used to assess job proficiency. Thorndike (12) has provided an extensive analysis and discussion of the varieties of measures that can be used for this purpose. An outline of his classification of measures is given:

- Evaluation of specific defined tasks
  - Evaluation of knowledge and understanding
  - Appraisals of actual job performance
  - Objective performance score
  - Observer scored job samples
  - Rated job samples
- Summary evaluations
  - Summary performance records
  - Summary academic grades
  - Summary ratings

Evaluation of defined tasks refers to the assessment of a limited segment of job behavior. It includes evaluation of knowledge and understanding as accomplished in paper-and-pencil job knowledge tests (the job knowledge tests used in Work Unit UTILITY fall into this class); objective performance scores based upon a permanent job record or product, such as a sample of typing or a shot pattern on a rifle target; observer-scored job samples, such as a performance check list in which steps in a job task or sequence are scored as having been performed or not (the job sample tests used in Work Unit UTILITY fall into this class); and rated job samples in which the adequacy of the performance of an entire task is rated by an observer. Summary evaluations include summary performance records such as the weekly output for piece workers or the annual sales record of a salesman; summary academic grades, and summary ratings such as a supervisor or peer rating of an employee (the supervisor ratings used in Work Unit UTILITY fall in this class).

While a discussion of the characteristics, advantages, and disadvantages of the different types of measures is beyond the scope of the present report (see Thorndike, 12, for an excellent treatment), it is readily seen that the suitability of a particular type of measure depends partly on the purpose for which information about performance is being gathered. Summary evaluations are inadequate for providing diagnostic information about a student’s strengths and weaknesses during training; job knowledge tests cannot provide information about a job incumbent’s skill, motivation, or productivity, and so on.

1 Published personality tests typically show a low relationship between measures of personality constructs and job performance. This is due, in part, to the general nature of such traits, as measured. Ratings of the specific behavioral characteristics required for performance in particular jobs would appear to provide a more promising approach to the measurement of personality traits.
The measurement of job proficiency can serve a variety of purposes. First, proficiency testing may be undertaken to assess an individual's performance of a particular task or job (i.e., to provide an estimate of the effectiveness of job performance or likely performance). Such information can be used for a variety of purposes: to determine whether a trainee is eligible for graduation at the end of training, for assigning him to a particular job or job level, and for establishing his suitability for advancement and promotion within a job.

Measurement of performance effectiveness requires assessment of the components of job performance—knowledge, skill, motivation, and personality characteristics. To evaluate knowledge, a job knowledge test is most appropriate. While knowledge can be evaluated by means of job sample tests, it is generally extremely inefficient to do so. However, where the knowledge that supports performance has not been adequately identified, appraisal of actual job performance, such as job sample tests, must be used. To evaluate skill, one of the varieties of job sample measures should be used. Objective performance scores or job sample scores are to be preferred to the more subjective, rated job samples. To evaluate motivation and personality characteristics, a summary rating is necessary. As indicated, published personality tests have generally proven unsatisfactory for this purpose. Rating scales for assessing motivation should be tied to specific observable job behaviors rather than being based upon poorly defined general impressions.

Second, proficiency testing may be undertaken to diagnose performance deficiencies either during training or on the job. Such testing has a pedagogical purpose only—to determine whether the trainee has or has not acquired specific knowledge and skill necessary to support performance and to permit the adjustment of training to rectify deficiencies.

To diagnose knowledge deficiencies, job knowledge tests are most appropriate. For this purpose, it is usually more efficient and informative to reduce all knowledge requirements to their simplest components. A series of multiple-choice items measuring the separate knowledges required for the accomplishment of a complex task is to be preferred to a single item that attempts to present simultaneously the full variety of stimulus possibilities the trainee would face in the performance of the actual job. Such an item provides only a global measure of the availability of many knowledges.

To diagnose skill deficiencies, observer-scored job samples, such as job sample tests, are most appropriate. Objective performance scores based upon a job record or product will usually not provide information in sufficient detail for diagnostic purposes. While rated job samples can be used, they are less preferred because of their subjective component.

Third, proficiency testing may be used to supplement the teaching process. The particular virtue of requiring a trainee to demonstrate information and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation. Individuals who know that they will be tested may be expected to concentrate on learning materials and skill in a test lies in the strong motivating properties of the testing situation.

Fourth, proficiency testing may be undertaken to evaluate the effectiveness of a training program. The development of a training program logically begins with the specification of certain performance capabilities that trainees are to acquire. Training content (skills and knowledge to be acquired), method, and organization are then developed to operate collectively in providing trainees going through the program with the desired capabilities. When training development has been completed, graduates of the program are tested to determine the extent to which these goals have been met.
Once the adequacy of a training program has been established through such an evaluation process, individuals who successfully complete the various subparts of the training program can be presumed to possess the desired capabilities. While end-of-training testing may still be conducted, such testing is no longer essential to ensure adequacy of the training product except as it may be used to determine the continuing quality of the program. To establish the adequacy of a training program (i.e., program evaluation), one of the various methods for appraising job performance is essential. Again, objective performance scores or job sample tests are preferred over rated job samples.

Fifth, proficiency testing may be undertaken to provide a criterion measure in research. Thus, the testing in Work Unit UTILITY was undertaken primarily to determine the job proficiency of men at different ability (AFQT) levels. Another frequent research purpose for proficiency tests is to validate other measures that are more remote from the ultimate criterion (actual job performance), but may be more convenient, economical, and suitable for operational use (13).

In addition to estimating the job proficiency of men in different ability groupings, the data obtained in UTILITY may be examined to provide information about the empirical relationships among the three criterion measures.

**DESCRIPTION OF THE UTILITY TESTS**

Job sample tests were used as the closest approximation possible to the ultimate criterion of actual job performance. Each job sample test comprised a number of subtests, was individually administered, and took from three and one-half to five hours to complete. To make testing realistic and to simulate job conditions with as much fidelity as possible, each subtest consisted of the performance of a single entire task with a natural beginning and ending. Tasks were composed of a series of actions or steps that would ordinarily be performed as part of a single operation in the performance of a job.

Subtest tasks or problems were introduced to subjects just as they might ordinarily encounter them. For example, one of the Repairman tasks involved the diagnosis and correction of a leaking oil seal in a fan shaft of a tank. The tank was prepared so that the oil seals leaked and the subject was told that oil was being thrown out of the tank’s grill doors. The Repairman was to locate the source of the oil leak and to repair the malfunction. In this problem, there were 19 necessary steps (Figure 1). Subjects earned a point for the correct performance of each step. This scoring procedure was used in all job sample tests.

The subtests contained in the four job sample tests are identified in Table 1.

In UTILITY, job knowledge (paper-and-pencil) tests were used to supplement job sample tests and provide an efficient means of sampling from the entire range of knowledge required in job performance. These tests, using a multiple-choice format, were similar to those traditionally used to assess learning in classroom situations and in Army training. However, in the development of these tests, particular care was taken to ensure that only the knowledge needed to support job performance was covered. Job-related information that is not required to support job performance has often been used to provide better discrimination among job incumbents, but it is invalid if the purpose of the test is to estimate job proficiency. A few illustrative items, taken from the Supply Specialist job knowledge test, are given in Figure 2.

Supervisor ratings were obtained for assessing motivation and other personal attributes, which the job incumbent had characteristically demonstrated in the performance of his job. This instrument consisted of the 14 rating scales that make up the Enlisted Efficiency Report used operationally throughout the Army. In Figure 3, two examples from the Supervisor Rating Scales are presented.
Steps in Fan Vertical Shaft Oil Seal Replacement Subtest—Repairman Job Sample Test

1. Testee removes a fan vane by removing four screws securing vane to housing.
2. Removes cotter pin, slotted hex nut, and flat washer, securing fan assembly to fan drive shaft.
3. Removes fan assembly from shaft.
4. Reports fan tower seals leaking.
5. Cuts and removes locking wire from bolts in oil seal housing.
6. Removes six capscrews securing housing.
7. Screws two 5/16-18 X 4 bolts in threaded holes in oil seal housing.
8. Turns both bolts in evenly until oil seal housing separates from fan drive housing cover.
9. Removes bolts from oil seal housing.

(Tester says: “FOR THE PURPOSE OF THIS TEST; WE WILL NOW CONSIDER THAT YOU HAVE REPLACED BOTH DEFECTIVE SEALS. HOW WOULD YOU REPLACE THE SEALS?”)

10. Testee answers “Remove old seal, clean housing, coat housing with adhesive cement, and install new seal.”

(Tester says: “NOW, YOU MUST PUT THIS VEHICLE IN SERVICEABLE CONDITION.”)

11. Secures oil seal housing with six bolts and flat washers.
12. Installs locking wire through bolt heads.
13. Replaces fan assembly on splined drive shaft.
14. Replaces flat washer on shaft.
15. Replaces slotted hex nut on shaft.
16. Torques nut to 50 lb-ft torque and aligns slot with hole in shaft.
17. Inserts new cotter pin through nut and shaft.
18. Checks clearance between each fan blade and fan housing with 0.070-inch feeler gauge. Adjusts fan housing as necessary.
19. Installs unmarked fan vane housing over rear fan or installs fan vane housing marked “Damper End” over front fan.

---

Figure 1

Table 1

<table>
<thead>
<tr>
<th>Title and Subtest Number</th>
<th>Description</th>
<th>Number of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armor Crewman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Disassembly and Assembly of the Cal. 45 Pistol, M1811A1</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>Disassembly and Assembly of the Cal. 45 Submachine Gun, M3A1</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Disassembly and Assembly of the Coaxial Machine Gun, M73</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>Disassembly and Assembly of the Cal. 50 Machine Gun, M85</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>Identification of Arm and Hand Flag and Light Signals</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>Identification of Combat Formations for Tanks</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>Mounting and Operating the AN/VRC 53 Tank Radio</td>
<td>10</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

**Job Sample Test**

<table>
<thead>
<tr>
<th>Title and Subtest Number</th>
<th>Description</th>
<th>Number of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Armor Crewman</strong> (Cont.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Identifying and Stating the Functions of the M60A1 Tank Controls</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>Starting and Stopping the Main Engine in Moderate, Hot, and Cold Weather</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>Driving the M60A1 Tank in Response to Arm and Hand Signals</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>Before Operations Maintenance of the M60A1 Tank</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>Mounting and Operating the AN/VRC 12 Tank Radio</td>
<td>14</td>
</tr>
<tr>
<td>13</td>
<td>Using the M1A1 Gunner’s Quadrant and Adjusting the Elevation Quadrant, M13A3</td>
<td>15</td>
</tr>
<tr>
<td>14</td>
<td>Preparation of Range Card</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>Loading the 105-mm Gun, M68</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>Reading and Setting the Ballistic Computer</td>
<td>22</td>
</tr>
<tr>
<td>17</td>
<td>Boresighting and Zeroing the 105-mm Gun</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>Checking the Azimuth Indicator for Accuracy and Slippage, and Putting the Turret Into Power</td>
<td>20</td>
</tr>
<tr>
<td>19</td>
<td>Adjustment of Fire in Response to Fire Commands</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>Reading and Setting the Azimuth Indicator</td>
<td>6</td>
</tr>
<tr>
<td><strong>Repairman</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Transmission Shift Linkage Adjustment (M60A1 Tank)</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>Transmission Servo Band Adjustment (M60A1 Tank)</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Voltage Regulator Adjusting Rheostat Adjustment (M60A1 Tank)</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Hydraulic Brake Pedal Adjustment (M60A1 Tank)</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>Fan Vertical Drive Shaft Oil Seal Replacement (M60A1 Tank)</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>Faulty Spark Plug Check (M151 Truck)</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>Ignition Timing Adjustment (M151 Truck)</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>Cylinder Compression Check (M151 Truck)</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Battery Hydrometer Check (M151 Truck)</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>Battery Condition Test (M151 Truck)</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>Front Outer Wheelbearing Adjustment</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>Fuel Pump Pressure Check (M35A1/A2 or M49C Truck)</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>Brake Bleeding (M35A1/A2 or M49C Truck)</td>
<td>17</td>
</tr>
<tr>
<td><strong>Supply Specialist</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Preparation of Laundry Roster</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>Preparation of Personal Clothing Request</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>Preparation of Hand Receipts</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>Sorting Equipment for Serviceability</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Preparation of Maintenance Request</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>Preparation of Statement of Charges</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>Preparation of Request for Issue</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>Maintenance of Document Register</td>
<td>15</td>
</tr>
</tbody>
</table>
Table 1 (Continued)

Job Sample Test

<table>
<thead>
<tr>
<th>Title and Subtest Number</th>
<th>Description</th>
<th>Number of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lighting the Fire Unit of the M1937 Field Range</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Preparation of Cocoa with Reconstituted Milk</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>Preparation of Scrambled Eggs</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Preparation of Jelly Roll</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>Preparation of Cook’s Worksheet</td>
<td>36</td>
</tr>
</tbody>
</table>

Illustrative Items From the Supply Specialist Job Knowledge Test

6. What is a Document Register used for?
   A. Filing hand receipts
   B. Issuing documents
   C. Recording document numbers
   D. Recording unserviceable equipment

12. What form is used to request single line item expendable supplies?
    A. DA Form 2765-1
    B. DA Form 2765
    C. DA Form 3161
    D. DA Form 2064

20. What regulation governs the issue and sale of personal clothing?
    A. AR 735-502-3
    B. AR 700-8400-1
    C. AR 712-320-5
    D. AR 725-220-3

38. What step is taken first when you receive the Equipment Status Report?
    A. The Company Commander signs the report
    B. You edit
    C. Increase the on-hand quantity
    D. You make an issue

66. You want to check to see if the forms you are using are current. What DA pamphlet index would you refer to?
    A. 310-1
    B. 310-2
    C. 310-3
    D. 310-4

88. When a recorded serial numbered item has been turned in how is it posted on the back of the Property Book page?
    A. Entered as a turn-in
    B. Lined out and initialed
    C. Debit posted and initialed
    D. Posted as a gain
Illustrative Items From the Supervisor Rating Scales
(Enlisted Efficiency Report)

ACCEPTABILITY (Ability to get along well with others)

a. _______ Never gets along well with other people
b. _______ Sometimes does not fit into group
c. _______ Gets along fairly well with others
d. _______ Gets along well with most people
e. _______ Gets along very well with others
f. _______ Gets along exceptionally well with others

RELIABILITY (Dependability and consistency of performance)

a. _______ Is not dependable
b. _______ Is sometimes late in finishing jobs
c. _______ Is average in reliability
d. _______ Works steadily
e. _______ Can always be relied upon to stick to the job
f. _______ Shows exceptionally dependable performance

Figure 3

In addition to the two examples given in Figure 3, the following characteristics were rated:

(1) Conduct (ability to control and manage personal behavior).
(2) Job performance (achievement of job tasks).
(3) Cooperativeness (ability to work effectively with others).
(4) Drive (energy devoted to job mastery and self-improvement).
(5) Development (effort directed toward realization of potential).
(6) Initiative (aggressive pursuit of methods to improve performance or productivity).
(7) Job knowledge (effort directed toward improvement of job knowledge).
(8) Application (willingness to work).
(9) Participation (contribution to group productive effort).
(10) Adaptability (ability to perform effectively in the face of changing job demands).
(11) Responsibility (willingness to accept responsibility).
(12) Leadership (ability to direct the work of others).

AN ANALYSIS OF THE INTERCHANGEABILITY OF PROFICIENCY MEASURES AND EVIDENCE FROM THE UTILITY DATA

When proficiency testing is undertaken to estimate the effectiveness of job performance, a combination of knowledge, skill, and motivational testing is often required. Nevertheless, practical constraints can be expected to promote the use of only those criterion measures that are simplest and most economical to administer. It is essential, therefore, to determine where one criterion instrument can reasonably be substituted for another.

1 Detailed descriptions of each instrument and its development and administration are provided in the second report of this series (2).
An instrument that has been designed to assess job motivation and other personal qualities, such as rating scales, is always essential if a complete estimate of job performance is desired. These results would be examined in conjunction with more direct measures of job skill and job knowledge. The decision as to the appropriateness of using a job sample test alone, a job knowledge test alone, or both, depends upon the weight of the skill and knowledge components in a job; this can be determined only after a careful analysis and specification of the information and skill needed to support performance.

Where job performance is mediated almost solely by knowledge per se, job knowledge tests should be used to assess the information a man has about his job. Job sample tests should not be used in such a situation since they are far more expensive to administer.

Where job performance relies almost solely on skill, job sample tests, or some other variety of performance measure, are essential. A convenient means for deciding on the skill requirements of a job, and hence the extent to which performance testing is necessary, is to determine how much practice or rehearsal is typically needed during training.¹

Where job performance depends upon a mix of knowledge and skill, both knowledge tests and performance tests are indicated. Thus, to determine whether a person should be given an automobile driver’s license, most states administer both a paper-and-pencil (job knowledge) test to assess knowledge about driving and a performance (job sample) test to assess driving skill or ability. Both types of test are necessary because each alone will not provide sufficient information about a license applicant’s capability to warrant issuing him a license. The knowledge test cannot reveal whether he has skill enough to drive an automobile, and the driving test cannot, without great expense, be made extensive enough to reveal his knowledge of driving rules, road signs, laws regulating vehicles, and so forth.

On the other hand, the ability of a clerk to perform his job is almost entirely dependent upon his possession of certain knowledge and information. The adequacy with which he makes entries in documents, and files material, can be established readily through knowledge testing (14). A job sample test is not necessary unless speed and neatness are critical; then a job sample test would also be necessary.

Job sample testing is expensive in time, money, development, and test administrators; therefore, reliance on job knowledge tests (and ratings of motivation and other personal characteristics) to assess job proficiency is desirable, if valid measurement can be accomplished in this manner. Yet, the correlations between job knowledge test scores and job sample criterion have often been found to be too low to support the use of job knowledge tests alone for assessing proficiency (13, 15, 16, 17, 18). Nevertheless, job knowledge tests should be valid for measuring proficiency in all jobs where skill components are minimal, if job knowledge tests are carefully constructed to measure only information that is directly relevant for performing the job. Job knowledge tests frequently correlate poorly with performance criteria because they include information that is irrelevant to job performance. For example, some varieties of theoretical information, terminology, and historical information, commonly found in training or course content, are readily available for the construction of test items. Because such items are often easy

¹It should be noted that a distinction has been made between skill that requires practice or rehearsal during training and skill that the trainee possesses at the outset of training and for which practice and rehearsal are not necessary. For example, there are skills that males usually acquire while growing up in our culture, such as the use of simple tools (screw driver, wrench, hammer, etc.). In general, such skill requirement can be overlooked as far as testing is concerned, if the trainee or job incumbent population can be assumed to possess the skill. However, if the population were to change, such an assumption would have to be reconsidered.
to prepare, they are frequently overrepresented in tests even though they do not bear directly upon performance of the job.

An analysis of the job sample tests used to measure performance in the four jobs studied in UTILITY showed that skill requirements in these jobs appear to be minimal. Table 2 shows the number of steps in each job sample test that were judged to be mediated by knowledge alone, and those judged to require cognitive skill or some combination of perceptual and motor skill. None of the steps in any of the performance tests required social skill. Practically all of the performance steps in each of the jobs can be supported by knowledge alone, and as was anticipated, skill requirements in the Supply Specialist job are virtually nonexistent.

This analysis suggests that the correlation between job sample test scores and job knowledge test scores in each of these jobs would be high, and, in all likelihood, the correlation would be highest in the Supply Specialist job.

Table 2

Frequency of Steps in Job Sample Tests Categorized According to Knowledge and Skill Requirements

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Armor Crewman</th>
<th>Repairman</th>
<th>Supply Specialist</th>
<th>Cook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Alone</td>
<td>338</td>
<td>165</td>
<td>153</td>
<td>145</td>
</tr>
<tr>
<td>Cognitive Skill and Knowledge</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Perceptual-Motor Skill and Knowledge</td>
<td>16</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total Number of Steps</td>
<td>354</td>
<td>175</td>
<td>156</td>
<td>158</td>
</tr>
</tbody>
</table>

INTERRELATIONSHIPS OF JOB SAMPLE TESTS, JOB KNOWLEDGE TESTS, AND SUPERVISOR RATINGS

Correlations among job sample scores, job knowledge scores, and supervisor ratings are given in Table 3. Correlations between these scores are presented for the combined Category IV and Non-IV groups. While the combined sample is artificial (i.e., not representative of the proportion of Category IVs to Non-IVs in the Army), the correlations for each group separately are so similar (Table 4) that the discussion of the implications of these correlations has been simplified by using the combined group correlation.

Here, as in an earlier report in this series (3), zero-order and partial correlations are both reported. The partial correlations provide indices of the strength of relationship between the criterion measures when the variance that each holds in common with job experience (months on the job) has been removed. Both the zero-order and partial

1 A few of the men tested in all the jobs held supervisory positions, undoubtedly requiring some degree of social skill. It has been anticipated however, that a test of supervisor skill would be inappropriate for the majority of the sample and a decision was made early in the research not to attempt to measure this aspect of performance.
Table 3
Correlations of Job Sample, Job Knowledge, and Supervisor Ratings for the Entire Sample

<table>
<thead>
<tr>
<th>Item</th>
<th>Armor Crewman (N=368)</th>
<th>Repairman (N=360)</th>
<th>Supply Specialist (N=380)</th>
<th>Cook (N=366)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero-Order Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Sample—Job Knowledge</td>
<td>.68</td>
<td>.59</td>
<td>.72</td>
<td>.58</td>
</tr>
<tr>
<td>Job Sample—Supervisor Ratings</td>
<td>.27</td>
<td>.20</td>
<td>.28</td>
<td>.28</td>
</tr>
<tr>
<td>Job Knowledge—Supervisor Ratings</td>
<td>.31</td>
<td>.23</td>
<td>.32</td>
<td>.35</td>
</tr>
<tr>
<td>Partial Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Sample—Job Knowledge</td>
<td>.49</td>
<td>.49</td>
<td>.65</td>
<td>.50</td>
</tr>
<tr>
<td>Job Sample—Supervisor Ratings</td>
<td>.13</td>
<td>.16</td>
<td>.24</td>
<td>.20</td>
</tr>
<tr>
<td>Job Knowledge—Supervisor Ratings</td>
<td>.21</td>
<td>.18</td>
<td>.29</td>
<td>.27</td>
</tr>
</tbody>
</table>

a All correlations are significantly different from zero (p<.05).

b Correlations with the effects of time on the job (MOJ) partialled out.

Table 4
Job Sample Test Reliabilities (Lower Limit)

<table>
<thead>
<tr>
<th>Job Title</th>
<th>(r_{\text{Cat-IV}})</th>
<th>(r_{\text{Non-IV}})</th>
<th>(r_{\text{Combined}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armor Crewman</td>
<td>.52</td>
<td>.55</td>
<td>.55</td>
</tr>
<tr>
<td>Repairman</td>
<td>.46</td>
<td>.38</td>
<td>.42</td>
</tr>
<tr>
<td>Supply Specialist</td>
<td>.51</td>
<td>.58</td>
<td>.54</td>
</tr>
<tr>
<td>Cook</td>
<td>.36</td>
<td>.34</td>
<td>.35</td>
</tr>
</tbody>
</table>

correlations show a high degree of relationship between job sample and job knowledge test scores, but a considerably lower degree of relationship between either of these criteria and supervisor ratings.

The correlations obtained here between job sample and job knowledge test scores are higher than those reported elsewhere (13) and tend to support the authors' hypotheses that job knowledge tests can be appropriately substituted for job sample tests, when a job contains little or no skill components and when only knowledge required on the job is used in the test.

During the preceding analysis of the job sample test step requirements (Table 2), the skill components in these four jobs were judged to be minimal. Even in the Repairman and Cook jobs, where manual operations and the use of tools are common, skill requirements appeared to be small. It may be assumed that most persons are able to perform in either of these jobs almost completely on the basis of knowledge of the required tasks and of the specific operations entailed. The magnitude of the correlations between job sample and job knowledge scores in all the jobs support this analysis. As anticipated, the highest correlation occurred between job sample and job knowledge...
criteria in the Supply Specialist job. This relationship is of particular significance for this analysis. A clerk's job, typified by the Supply Specialist, represents one of the purest examples of a job where knowledge rather than skill is sufficient to support performance.

If job knowledge tests are to be used to estimate job proficiency, the reliability of these measures must also be considered. Reliability of job knowledge tests was obtained using an odd-even split of test items. Reliability coefficients, for both Category IVs and Non-IVs in each MOS are presented in Table 5. Both the computed reliability based upon the odd-even halves and the reliabilities corrected for test length (Spearman-Brown prophecy formula) are given.

<table>
<thead>
<tr>
<th>Job Title</th>
<th>Category</th>
<th>Computed r</th>
<th>r Corrected for Test Length</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armor Crewman</td>
<td>Cat-IV</td>
<td>.70</td>
<td>.82</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Non-IV</td>
<td>.65</td>
<td>.79</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repairman</td>
<td>Cat-IV</td>
<td>.61</td>
<td>.76</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>Non-IV</td>
<td>.63</td>
<td>.77</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supply Specialist</td>
<td>Cat-IV</td>
<td>.88</td>
<td>.94</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>Non-IV</td>
<td>.80</td>
<td>.89</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cook</td>
<td>Cat-IV</td>
<td>.69</td>
<td>.82</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Non-IV</td>
<td>.74</td>
<td>.85</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The high reliabilities obtained for the job knowledge tests are easily sufficient to permit the application of such tests in situations where group measurement is required, as, for example, in the UTILITY research. Further, these reliabilities appear to be of sufficient magnitude to allow the tests to be used in the measurement of individuals.

The lower limits of the reliabilities of the job sample tests were estimated using the correlations between job sample and job knowledge test scores and the reliabilities computed for the job knowledge test scores

\[ r_{xy}^{\max} = \sqrt{(r_{xx})(r_{yy})} \]

and solving the equation for the unknown job sample reliability. These estimates of the lower limits of the job sample test reliabilities are presented in Table 4.

It was not appropriate to directly compute reliabilities for job sample tests for the following reasons: (a) Job sample subtests were heterogeneous (i.e., they measured different skills of varying degrees of complexity and difficulty). Under these conditions, a test-retest estimate of reliability is appropriate rather than some estimate of internal consistency. No retest data were possible in this study. (b) In general, within any subtest steps were not independent. (c) Not all subjects performed all steps (if a subject failed to complete three successive steps, the subtests were discontinued).
An examination of the correlations between supervisor ratings and the two test criteria (Table 3) shows relatively weak relationships; this is as expected. The supervisor ratings were designed to measure general behavioral characteristics such as adaptability, reliability, drive and conduct—a composite of attributes that underlie general work and interpersonal effectiveness (19). This instrument was included in UTILITY in order to obtain estimates of these attributes that were not obtained through job sample or job knowledge testing.

As seen, there is some degree of relationship between supervisor ratings and the other two criterion measures. Two of the 14 scales in the supervisor rating provide specific ratings of job knowledge and job performance, thereby making a direct though small contribution to the relationships. Also, most of the personal attributes would be expected to contribute to the acquisition of specific knowledge and skill. Thus, there is some amount of correlation between supervisor ratings and the other two criteria. Nevertheless, the overall relationship between supervisor ratings and the two test instruments are low, since the former is also measuring attributes other than job knowledge and job performance per se.¹

RELATIONSHIPS BETWEEN JOB SAMPLE TESTS AND JOB KNOWLEDGE TESTS IN SELECTED SUBSAMPLES

Having observed the relatively high correlations between job sample and job knowledge test scores for the entire UTILITY sample, it is appropriate to determine whether or not this relationship is maintained in particular subgroups. It is possible that two variables may correlate differently in a particular subgroup than in the overall sample. Were there to be a significantly lower correlation between job sample and job knowledge scores for a subgroup, it would be inappropriate to substitute job knowledge tests for job sample tests as a measure of job proficiency for that subgroup (without making adjustment to take such differences into account).

Scores on job knowledge tests tend to be highly related to reading ability (20). Therefore, it is appropriate to examine the correlations of the two test criteria in subgroups with low and high reading ability, and in subgroups that differ in characteristics related to general verbal ability to determine whether the relationships between job sample and job knowledge tests are maintained in these groups.² Category IVs, men with low verbal aptitude and men with low education, were all found to have significantly lower reading ability (20). Blacks in the sample also tended to have lower reading ability (21). Correlations between job sample and job knowledge scores for the entire sample are compared with those for men with low and high reading ability (Table 6), Category IVs and Non-IVs (Table 7), men with low and high verbal aptitude (Table 8), men with low and high education levels (Table 9), and Blacks and Whites (Table 10).

¹The conditions under which the UTILITY study was conducted did not permit inter-rater reliability to be obtained for the supervisor ratings. Reliabilities based upon at least two raters would provide the most meaningful measures of stability of these instruments. Estimates of the lower limit of supervisor rating reliabilities, obtained in the same manner as those computed for job sample tests, would be inappropriate and expected to provide gross underestimates of reliability.

²Reading ability was assessed using the California Test Bureau Survey of Reading Achievement, Junior High Level (grades 7-9); verbal ability was assessed using the Verbal Test of the Army Classification Battery; educational level was obtained from a questionnaire item: Years of Education Completed. Low and high groups were obtained for each of these variables by assigning people to groups on or above the median and below the median.
Table 6

Correlations of Job Sample and Job Knowledge for Low Reading and High Reading Ability Subgroups

<table>
<thead>
<tr>
<th>Sample</th>
<th>Armor Crewman</th>
<th>Repairman</th>
<th>Supply Specialist</th>
<th>Cook</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>r</td>
<td>N</td>
<td>r</td>
</tr>
<tr>
<td>Zero-Order Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .68</td>
<td>360 .59</td>
<td>380 .72</td>
<td>366 .58</td>
</tr>
<tr>
<td>Low Reading</td>
<td>178 .65</td>
<td>174 .45</td>
<td>186 .67</td>
<td>176 .42</td>
</tr>
<tr>
<td>High Reading</td>
<td>190 .57</td>
<td>186 .47</td>
<td>194 .64</td>
<td>190 .54</td>
</tr>
<tr>
<td>Partial Correlations C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .49</td>
<td>360 .49</td>
<td>380 .65</td>
<td>366 .50</td>
</tr>
<tr>
<td>Low Reading</td>
<td>178 .34</td>
<td>174 .44</td>
<td>186 .59</td>
<td>176 .49</td>
</tr>
<tr>
<td>High Reading</td>
<td>190 .51</td>
<td>186 .48</td>
<td>194 .64</td>
<td>190 .41</td>
</tr>
</tbody>
</table>

All correlations are significantly different from zero (p<.05).

Ns for subsamples will occasionally not add to total for entire sample.

Where pertinent data were missing in Army records, a man was omitted from an analysis.

Correlations with the effects of MOJ partialled out.

Table 7

Correlations of Job Sample and Job Knowledge for Category IV and Non-Category IV Subgroups

<table>
<thead>
<tr>
<th>Sample</th>
<th>Armor Crewman</th>
<th>Repairman</th>
<th>Supply Specialist</th>
<th>Cook</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>r</td>
<td>N</td>
<td>r</td>
</tr>
<tr>
<td>Zero-Order Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .68</td>
<td>360 .59</td>
<td>380 .72</td>
<td>366 .58</td>
</tr>
<tr>
<td>Category IV</td>
<td>186 .65</td>
<td>178 .59</td>
<td>188 .69</td>
<td>186 .54</td>
</tr>
<tr>
<td>Non-Category IV</td>
<td>182 .66</td>
<td>182 .54</td>
<td>192 .72</td>
<td>180 .54</td>
</tr>
<tr>
<td>Partial Correlations C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .49</td>
<td>360 .49</td>
<td>380 .65</td>
<td>366 .50</td>
</tr>
<tr>
<td>Category IV</td>
<td>186 .34</td>
<td>178 .47</td>
<td>188 .58</td>
<td>186 .42</td>
</tr>
<tr>
<td>Non-Category IV</td>
<td>182 .49</td>
<td>182 .45</td>
<td>192 .64</td>
<td>180 .45</td>
</tr>
</tbody>
</table>

All correlations are significantly different from zero (p<.05).

Ns for subsamples will occasionally not add to total for entire sample.

Where pertinent data were missing in Army records, a man was omitted from an analysis.

Correlations with the effects of MOJ partialled out.
### Table 8

**Correlations of Job Sample and Job Knowledge for Low Verbal and High Verbal Aptitude Subgroups**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Armor Crewman</th>
<th>Repairman</th>
<th>Supply Specialist</th>
<th>Cook</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N&lt;sup&gt;b&lt;/sup&gt;</td>
<td>r</td>
<td>N</td>
<td>r</td>
</tr>
<tr>
<td>Zero-Order Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .68</td>
<td>360 .59</td>
<td>380 .72</td>
<td>366 .58</td>
</tr>
<tr>
<td>Low Verbal</td>
<td>178 .66</td>
<td>174 .55</td>
<td>174 .68</td>
<td>174 .53</td>
</tr>
<tr>
<td>High Verbal</td>
<td>181 .68</td>
<td>179 .55</td>
<td>179 .72</td>
<td>186 .59</td>
</tr>
<tr>
<td>Partial Correlations&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .49</td>
<td>360 .49</td>
<td>380 .65</td>
<td>366 .50</td>
</tr>
<tr>
<td>Low Verbal</td>
<td>178 .40</td>
<td>174 .44</td>
<td>174 .58</td>
<td>174 .41</td>
</tr>
<tr>
<td>High Verbal</td>
<td>181 .53</td>
<td>179 .48</td>
<td>179 .65</td>
<td>186 .54</td>
</tr>
</tbody>
</table>

<sup>a</sup> All correlations are significantly different from zero ($p<.05$).

<sup>b</sup> Ns for subsamples will occasionally not add to total for entire sample.

Where pertinent data were missing in Army records, a man was omitted from an analysis.

<sup>c</sup> Correlations with the effects of MOJ partialled out.

### Table 9

**Correlations of Job Sample and Job Knowledge for Low Education and High Education Level Subgroups**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Armor Crewman</th>
<th>Repairman</th>
<th>Supply Specialist</th>
<th>Cook</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N&lt;sup&gt;b&lt;/sup&gt;</td>
<td>r</td>
<td>N</td>
<td>r</td>
</tr>
<tr>
<td>Zero-Order Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .68</td>
<td>360 .59</td>
<td>380 .72</td>
<td>366 .58</td>
</tr>
<tr>
<td>Low Education</td>
<td>142 .68</td>
<td>150 .63</td>
<td>99   .77</td>
<td>140 .50</td>
</tr>
<tr>
<td>High Education</td>
<td>222 .67</td>
<td>209 .56</td>
<td>281 .70</td>
<td>223 .62</td>
</tr>
<tr>
<td>Partial Correlations&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368 .49</td>
<td>360 .49</td>
<td>380 .65</td>
<td>366 .50</td>
</tr>
<tr>
<td>Low Education</td>
<td>142 .40</td>
<td>150 .44</td>
<td>99   .58</td>
<td>140 .41</td>
</tr>
<tr>
<td>High Education</td>
<td>222 .53</td>
<td>209 .48</td>
<td>281 .65</td>
<td>223 .54</td>
</tr>
</tbody>
</table>

<sup>a</sup> All correlations are significantly different from zero ($p<.05$).

<sup>b</sup> Ns for subsamples will occasionally not add to total for entire sample.

Where pertinent data were missing in Army records, a man was omitted from an analysis.

<sup>c</sup> Correlations with the effects of MOJ partialled out.
Table 10

Correlations of Job Sample and Job Knowledge for Black and White Subgroups

<table>
<thead>
<tr>
<th>Sample</th>
<th>Armor Crewman</th>
<th>Repairman</th>
<th>Supply Specialist</th>
<th>Cook</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>r</td>
<td>N</td>
<td>r</td>
</tr>
<tr>
<td>Zero-Order Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368</td>
<td>.68</td>
<td>360</td>
<td>.59</td>
</tr>
<tr>
<td>Black</td>
<td>70</td>
<td>.80</td>
<td>46</td>
<td>.55</td>
</tr>
<tr>
<td>White</td>
<td>294</td>
<td>.65</td>
<td>312</td>
<td>.59</td>
</tr>
<tr>
<td>Partial Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Sample</td>
<td>368</td>
<td>.49</td>
<td>360</td>
<td>.49</td>
</tr>
<tr>
<td>Black</td>
<td>70</td>
<td>.48</td>
<td>46</td>
<td>.30</td>
</tr>
<tr>
<td>White</td>
<td>294</td>
<td>.46</td>
<td>312</td>
<td>.51</td>
</tr>
</tbody>
</table>

All correlations are significantly different from zero (p<.05).

Ns for subsamples will occasionally not add to total for entire sample.

Where pertinent data were missing in Army records, a man was omitted from an analysis.

Correlations with the effects of MOJ partialled out.

In each Army job studied, the zero-order correlations for the subgroups are similar to each other. While the partial correlations (with months of job experience partialled out) occasionally show somewhat greater separation, suggesting some possible effect of restriction in range on job knowledge test scores in groups with lower reading ability, only one (out of 20) of the differences between pairs of correlations for the subsamples is statistically significant. This difference was found in Armor Crewman between the low and high reading groups.

CONCLUSION

On the basis of the relatively low skill requirements identified in the analysis of job sample test steps and the high degree of correlation between job sample test scores and job knowledge test scores, it appears reasonable to conclude that in the four occupational specialties studied, the job knowledge tests could appropriately be used in lieu of the job sample tests.

As indicated earlier, job knowledge tests can be used instead of job sample tests in occupational specialties if skill components in such jobs are minimal.

In the development of job knowledge tests, it is crucial that the tests require only knowledge that is needed in the actual performance of the job. Guidance for the preparation of written test items should be given careful and rigorous attention (22).

The estimates of the lower limit of reliability for the job sample tests indicate that these tests are adequate where group measurement is undertaken; no conclusion regarding adequacy of reliability for individual measurement can be made. The high reliabilities of the job knowledge tests allow their use where group or individual measurement is desired.
Both the job knowledge tests and the job sample tests possessed reliability characteristics making them suitable for use in the UTILITY research. Since only lower limit reliability estimates are available, it is not possible to discuss comparisons between the job knowledge and job sample reliabilities.¹

¹It is likely that the reliabilities reported for both the job knowledge and job sample tests represent underestimates of the degree to which they are free of errors of measurement, since tests were heterogeneous. (See earlier discussion of the composition of the job sample subtests and the examples of the job knowledge test items.) Such an underestimate is particularly likely in the case of the Cook job sample test where the subtests are dissimilar, ranging across fueling and lighting a field range, preparing food, and planning the preparation of a meal, including assignment of kitchen duties, timing the cooking sequence, and use of left-overs (using the Cook’s Worksheet).
LITERATURE CITED


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

President
Executive Vice President
Dr. Meredith P. Crawford
Dr. William A. McClelland

Director for Business Affairs and Treasurer
Mr. Charles W. Smith
Mr. Arnold A. Heyl

Director for Operations
Dr. Robert G. Smith, Jr.

Director for Program Development
Dr. Eugene A. Cogan

Director for Research Design and Reporting

RESEARCH DIVISIONS

HumRRO Division No. 1 (System Operations)
300 North Washington Street
Alexandria, Virginia 22314

HumRRO Division No. 2
Fort Knox, Kentucky 40121

HumRRO Division No. 3
Post Office Box 5787
Presidio of Monterey, California 93940

HumRRO Division No. 4
Post Office Box 2066
Fort Benning, Georgia 31905

HumRRO Division No. 5
Post Office Box 6057
Fort Bliss, Texas 79916

HumRRO Division No. 6 (Aviation)
Post Office Box 428
Fort Rucker, Alabama 36360

HumRRO Division No. 7 (Social Science)
300 North Washington Street
Alexandria, Virginia 22314

Dr. J. Daniel Lyons
Director

Dr. Donald F. Haggard
Director

Dr. Howard H. McFann
Director

Dr. T.G. Jacobs
Director

Dr. Albert L. Kubala
Director

Dr. Wallace W. Prophet
Director

Dr. Arthur J. Hoehn
Director