

AD-A182 387

DISTRIBUTION OF QUALITY PROGRAM HANDBOOK(U) ARMY
SOLDIER SUPPORT CENTER-NATIONAL CAPITAL REGION
ALEXANDRIA VA JUN 87

1/2

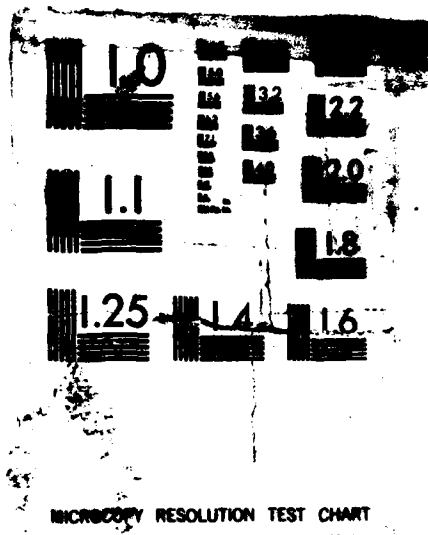
UNCLASSIFIED

F/G 5/9

NL

cont.



MICROCOPY RESOLUTION TEST CHART

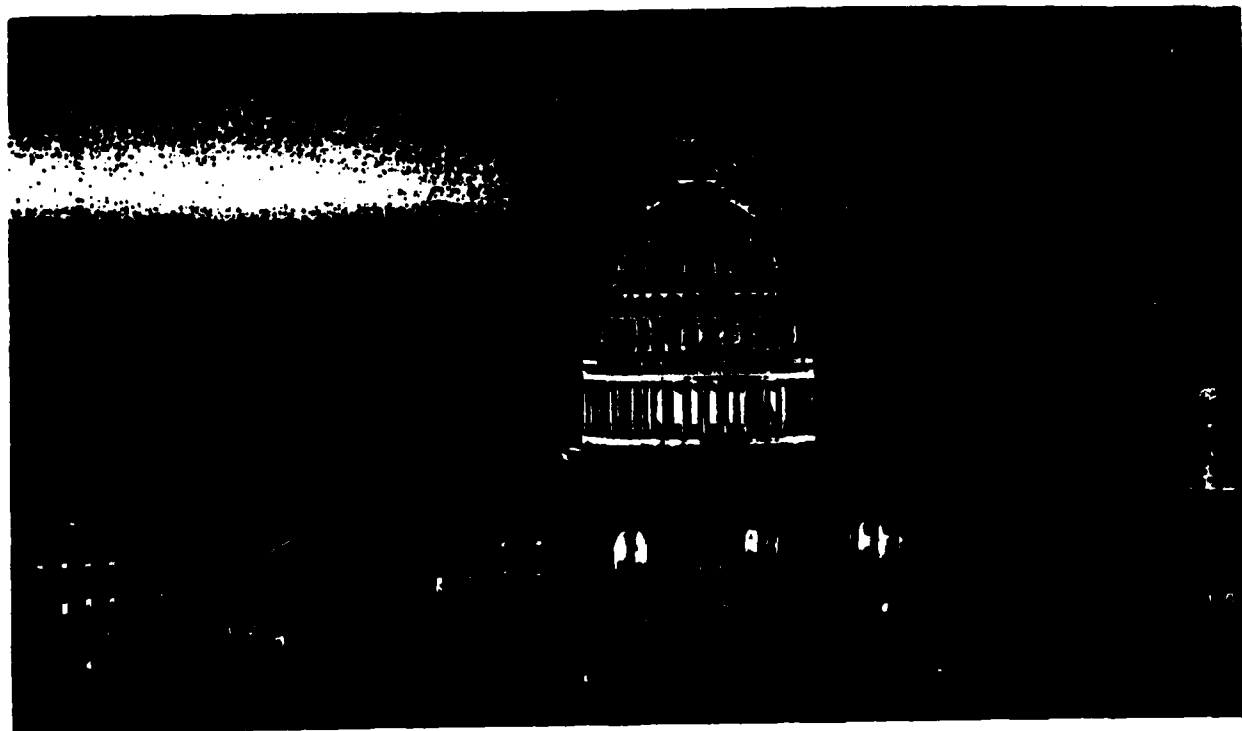
1

JUNE 1987

DTIC FILE COPY

DISTRIBUTION OF QUALITY PROGRAM HANDBOOK

AD-A182 307



SSC-NCR

DTIC
ELECTE
JUL 15 1987
S D
E

This document has been approved
for public release and sale; its
distribution is unlimited.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS NA	
2a. SECURITY CLASSIFICATION AUTHORITY NA			3. DISTRIBUTION/AVAILABILITY OF REPORT Distribution Statement A	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE				
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NA			5. MONITORING ORGANIZATION REPORT NUMBER(S) NA	
6a. NAME OF PERFORMING ORGANIZATION US Army Soldier Support Center-- National Capital Region		6b. OFFICE SYMBOL (if applicable) ATNC-NMF-A	7a. NAME OF MONITORING ORGANIZATION	
6c. ADDRESS (City, State, and ZIP Code) Alexandria, Virginia 22332-0400			7b. ADDRESS (City, State, and ZIP Code)	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION		8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State, and ZIP Code)			10. SOURCE OF FUNDING NUMBERS	
			PROGRAM ELEMENT NO.	PROJECT NO.
			TASK NO.	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Distribution of Quality Program Handbook (Unclassified)				
12. PERSONAL AUTHOR(S)				
13a. TYPE OF REPORT Information		13b. TIME COVERED FROM TO		14. DATE OF REPORT (Year, Month, Day) 8706
15. PAGE COUNT 56				
16. SUPPLEMENTARY NOTATION NA				
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP		
05	09			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) The <u>Distribution of Quality Program Handbook</u> explains what the Distribution of Quality (DQ) Program is, defines the specifications that proponent DQ submissions must meet, and describes one possible method to meet these specifications.				
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Leslie B. Lord			22b. TELEPHONE (Include Area Code) 202-325-3703	22c. OFFICE SYMBOL ATNC-NMF-A

DISTRIBUTION OF QUALITY (DQ)

PROGRAM HANDBOOK

JUNE 1987

The Distribution of Quality Handbook and the concepts herein were developed by the Personnel Programs and Analysis Branch, Personnel Plans Division, Manning Integration Directorate, of the United States Army Soldier Support Center--National Capital Region. Questions and comments concerning any concepts within this Handbook should be addressed to:

COMMANDER
USA Soldier Support Center--NCR
Attn: ATNC-MNF-A
200 Stovall Street
Alexandria, Virginia 22332-0400

TELEPHONE: COMM (703) 325-3703
AUTOVON 221-3703



Approved For	
By	
Date	
Initials	
Signature	
Comments	

A-1

TABLE OF CONTENTS

CHAPTER 1	4
OVERVIEW OF THE DISTRIBUTION OF QUALITY (DQ) PROGRAM .	4
What Is the DQ Program?	4
Why the DQ Program?	4
DQ and the Skill Level 3 (SL3) Data Program . . .	5
DQ and Project A	6
Primary Organizations and their Responsibilities .	7
CHAPTER 2	10
SPECIFICATIONS FOR PROPONENT DQ REQUIREMENTS	
SUBMISSIONS	10
Step 1: Identify the Target Skill Level	10
Step 2: Select the measure(s) of job performance .	10
Step 3: Validate the measure(s) of job	
performance	11
Step 4: Set the minimum acceptable score	11
Step 5: Administer the performance measure(s) and	
collect the data	11
Step 6: Quantify the relationship between quality	
and performance	11
Step 7: Calculate the Skill Level 1 accession	
Distribution of Quality mix required . .	11
Step 8: Assess the probable impact of projected	
changes on the required SL1 accession DQ	
mix	11
Step 9: Prepare and submit the analysis report .	12
CHAPTER 3	15
DQ PROCESS AND MILESTONES	15
APPENDIX A	17
ARMED FORCES QUALIFICATION TEST (AFQT) CATEGORY TABLE .	17
APPENDIX B	18
DETERMINING CONTENT VALIDITY	18
APPENDIX C	21
DETERMINING RELIABILITY	21
APPENDIX D	26
TESTING THE SIGNIFICANCE OF r	26
APPENDIX E	27
DETERMINING THE MINIMUM ACCEPTABLE SCORE	27
APPENDIX F	34
TESTING THE STATISTICAL SIGNIFICANCE OF THE DIFFERENCE	
BETWEEN TWO MEANS	34

APPENDIX G	38
ESTABLISHING DQ REQUIREMENTS	38
APPENDIX H	45
SAMPLE DQ ANALYSIS REPORT	45

CHAPTER 1

OVERVIEW OF THE DISTRIBUTION OF QUALITY (DQ) PROGRAM

What Is the DQ Program?

The Distribution of Quality (DQ) Program is the system Proponents use to justify the percentages of recruits in AFQT Categories I-III A, IIIB, and IV needed in their accession MOS each year. The primary considerations in this process are the recruits' probability of achieving successful entry level training/job performance and ensuring an adequate pool of high quality personnel from which to grow future NCO leadership. The working definition of a high quality recruit is a high school diploma graduate whose Armed Forces Qualification Test (AFQT) score percentile is in the top half of the national population (Categories I-III A). By law, the percentage of Category IV recruits cannot exceed 20 percent per year, and AFQT Category V personnel are not eligible for military service. A table showing the six AFQT categories and the percentile ranges of each is located in Appendix A.

As the proponent staff integrator for both the Department of the Army and the Training and Doctrine Command (TRADOC), the Soldier Support Center--National Capital Region (SSC-NCR) is responsible for developing and managing the Distribution of Quality Program. SSC-NCR's duties include acting as the TRADOC executive agent and the Proponents' advocate in working with the Office of the Army Deputy Chief of Staff for Personnel (ODCSPER) and the Recruiting Command (USAREC) to establish the DQ targets for each MOS in the development of the Army's annual recruiting mission.

Why the DQ Program?

The Distribution of Quality Program was instituted in 1982 to influence the placement of high quality recruits into the Army's various accession MOS. Minimum Aptitude Area (AA) score requirements for and the market appeals of each MOS had previously been the driving factors of recruit placement. Although AA scores predict the ability of soldiers to complete entry level training for an MOS, they do not predict soldiers' abilities to meet technical skill and leadership requirements beyond entry level training. Under the DQ Program, TRADOC provides a desired profile of recruits by AFQT category based on training needs and career force requirements in the out years for

every accession MOS. This profile enables USAREC to do a better job of distributing high quality recruits.

Since 1982, Congress has been closely examining USAREC's budget in conjunction with the Army's recruiting successes and has repeatedly asked why the Army needs to recruit so many high quality soldiers. The only way the Army can solidly justify its aggregate recruit quality needs to Congress is to develop such justification for every accession MOS. These justifications must be based on the relationship of successful job/training performance to AFQT score category.

The Army was enjoying unprecedented success in recruiting a high percentage of high quality recruits when the DQ Program was proposed and initially implemented. In this era of tightening resource constraints, it is essential that the Army effectively manage the MOS distribution of its high quality recruits; therefore, the mission of the DQ Program is now more important than ever.

TRADOC submitted proponent DQ requirements to ODCSPER and USAREC for FY84-87, but their influence was considerably diminished by the insufficient justifications given for most of the proponents' stated needs. It is essential that proponents use job/training performance measures as a basis for justifying their soldier accession quality requirements. Continued failure of most of the proponents to significantly improve their recruit quality justifications may result in lower percentages of high quality recruits in the future.

DQ and the Skill Level 3 (SL3) Data Program

SSC-NCR is TRADOC's executive agent for the Skill Level 3 (SL3) Data Program. The DQ and the SL3 Data Programs are parallel attempts to achieve the same purpose. The SL3 Data Program was implemented at the request of the Vice Chief of Staff of the Army (VCSA) to supplement the DQ Program. The DQ Program was designed to enable the Army to base its analyses of personnel quality requirements on the relationship of NCO quality to performance. The intent was for an annual test program to evaluate NCO performance and quality for some MOS each year until the personnel quality needs for all MOS are certified and to have whatever annual re-looks are needed in future years.

There have been two SL3 Data submissions to ODCSPER, but the data of both submissions have proved to be of little utility in supplementing DQ due to a number of program implementation problems. Proponents who choose to follow the procedures outlined in this handbook will be able to meet the requirements of both the DQ and the SL3 Data Programs simultaneously.

DQ and Project A

Project A is the Army's portion of a Joint Service project directed into existence by Congress and the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics) in 1980 to link enlistment standards to job performance. The Army Research Institute for the Behavioral and Social Sciences (ARI) is responsible for accomplishing this mission. The major thrusts of this project are to determine how well the Armed Services Vocational Aptitude Battery (ASVAB) predicts performance on the job and to evaluate other potential selection and classification measures. The 21 MOS included in this long term research project were chosen to be maximally representative of all entry level specialties. Before Project A, it was known that the current ASVAB is a valid predictor of entry level training success, but there was a need to establish its validity for predicting actual job performance and to develop new tests for a wider range of attributes that contribute to performance (e.g., motivation, psychomotor abilities, and spatial ability). A number of these attributes has been identified, and new tests of them are being validated for these 21 MOS. For this effort, ARI has developed a wide range of performance indicators, including hands-on tests (for 9 MOS), written tests, and ratings by peers and supervisors. Also, scores on the Skill Qualification Tests (SQTs) for more than 265,000 soldiers in 172 MOS are being used to validate the ASVAB. Methods for generalizing the results from the 21 Project A MOS to the domain of all entry level specialties are being developed. The product of these efforts will be a proposal to augment the ASVAB. This product will not set minimum standards for entry into training for any MOS, but it should enable the Army to improve its process of matching new soldiers to the MOS for which they are best qualified.

If the Army chooses to implement fully for all MOS such an expanded ASVAB, the DQ and the SL3 Data Programs in their present form will probably become extinct; however, it is highly improbable that such implementation could be accomplished prior to the mid-1990s. Meanwhile, the annual DQ mission must be accomplished using the tools currently available (i.e., proponent validated performance measures plus the AFQT and the appropriate AA scores from the current ASVAB) to effectively manage its distribution of high quality recruits and answer Congress' questions on its recruit accession quality needs. Any available Project A data that can help the DQ Program achieve these objectives should be used.

Primary Organizations and their Responsibilities

Success of the DQ Program is dependent upon the coordinated and cooperative efforts of many people in various parts of the Army. The primary organizations involved in this mission and their functions are:

1. Proponent Schools

- >Determine MOS performance requirements
- >Justify DQ goals based on MOS performance requirements

2. SSC-NCR

- >Provides guidance to Proponents
- >Provides technical assistance to Proponents
- >Collects, consolidates, and submits Proponents' DQ requirements
- >Analyzes Proponents' justifications and other factors to provide alternatives/recommendations for trade-off analysis decisions
- >Represents Proponents and TRADOC in trade-off analysis with USAREC and ODCSPER

3. Headquarters, TRADOC

- >Provides guidelines/sets priorities for DQ mix, if necessary
- >Participates in trade-off analysis with USAREC, ODCSPER, and SSC-NCR, if necessary

4. USAREC

- >Provides information to ODCSPER on what DQ accession goals are feasible based on market and resource projections
- >Recruits the DQ mix as set by ODCSPER in the annual recruiting mission

5. ODCSPER

- >Reviews TRADOC's DQ requirements and USAREC's DQ proposal based on economic feasibility

>Conducts trade-off analysis between TRADOC and USAREC

>Determines USAREC DQ production goals for annual mission

From this list, it is apparent that the Proponents' requirements are reviewed and analyzed several times before the recruiting mission is set. The DQ decision process involves interaction among SSC-NCR/TRADOC, ODCSPER, and USAREC. This process and the primary issues involved are shown in Figure 1.

DQ DECISION PROCESS

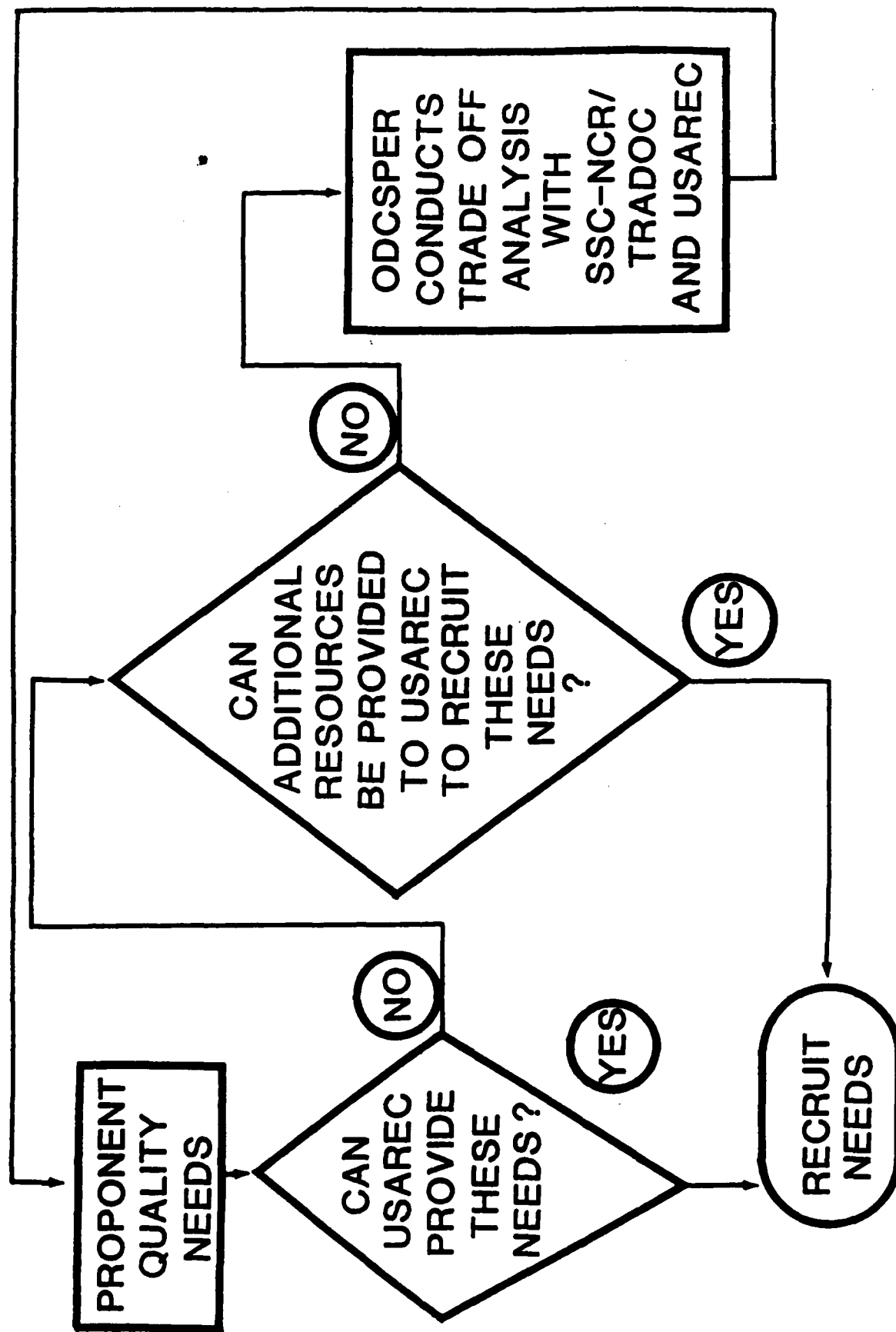


FIGURE 1

CHAPTER 2

SPECIFICATIONS FOR PROPONENT DQ REQUIREMENTS SUBMISSIONS

The process of determining Proponent accession quality requirements includes both objective and subjective factors; i.e., objective tools are used to assist in making subjective but credible assessments. In most schools, this process will probably be an integration of efforts by the Proponent Office, Directorate of Combat Development (DCD), the Directorate of Training and Doctrine (DOTD), the Directorate of Evaluation and Standardization (DOES), the training departments, and possibly some other offices.

In this chapter are the general procedures that Proponents must follow to justify their DQ requirements for each of their accession MOS and the minimum standards that each DQ submission must meet at every step to be considered credible. For some of these steps, there may be a number of different ways to accomplish them and still meet the requirements of these specifications. In such cases, appropriate appendices have been placed in this handbook to provide the implementation details of a suggested methodology for those Proponents who choose to use it. Proponents who choose to use some other methodology to meet these requirements must obtain prior approval of their methodologies from SSC-NCR. In addition to following the procedures outlined in this chapter, Proponents should also include in their assessments such factors as the impact of new equipment scheduled for deployment and projected changes in doctrine, training, and organization within the next five years.

Step 1: Identify the Target Skill Level. Target Skill Level is defined as that level which is most critical to mission success or failure in an MOS. Soldiers at this skill level have full technical proficiency to do the work and have earned the appropriate rank to direct/lead subordinates in daily job performance. They can also provide technical advice to the NCOs/officers above them. Traditional thinking is that Skill Level 3 meets this definition in most MOS; however, such may not be the case in many MOS. Proponents must make and justify this determination, and justifications must be especially strong if the Target Skill Level is other than SL3.

Step 2: Select the measure(s) of job performance. The performance measures must test enough critical tasks to be representative of the entire job of soldiers at the Target Skill Level of the MOS. Usually, proxies for measures of job performance are used. These proxies measure individual scores on job proficiency (hands-on tests) and/or job knowledge (written tests). Usable performance measures probably already exist in

many, if not most, MOS. Since the purpose here is to measure performance in a specific MOS, tasks from the Common Task Testing (CTT) Manual would not usually be included in these measures.

Step 3: Validate the measure(s) of job performance. The two issues of primary concern here are what a test measures (validity) and the quality of the measure (reliability). Validity is concerned with whether or not the performance measure is representative of the job and is usually a rational, judgmental process (see Appendix B). Reliability refers specifically to the consistency of the performance measure. The most widely used approach is the Kuder-Richardson method, which is a particularly appropriate measure of test consistency (see Appendices C and D).

Step 4: Set the minimum acceptable score on the measure(s) of job performance. This score should be the minimum level of performance for successful job accomplishment. If possible, it should be tied to the minimum level necessary for mission success. Reasons for selecting this minimum acceptable performance standard must also be included in the submission (see Appendix E).

Step 5: Administer the performance measure(s) and collect the data. A sufficient number of soldiers must be tested to permit statistically valid findings, and the soldiers selected for testing must be representative of the MOS population at the Target Skill Level. If appropriate, information on factors which may have affected performance (e.g., weather) should also be collected.

Step 6: Quantify the relationship between quality and performance. The mathematical relationship between the ASVAB scores and the performance of soldiers at the Target Skill Level must be determined. A step by step explanation of how to test the mathematical significance between ASVAB data and performance scores is located in Appendix F.

Step 7: Calculate the Skill Level 1 accession Distribution of Quality mix required to meet the quality requirements at the Target Skill Level. This calculation process is often called back aging. A methodology for back aging DQ data is explained in Appendix G.

Step 8: Assess the probable impact of projected changes on the required SL1 accession DQ mix. Accomplishing this step merely requires an estimate of how much the quality mix derived from following Steps 1-7 should be augmented or diminished due to projected changes in equipment, doctrine, training, and organization during the next five years. Although this process is subjective, cogent logic must be given to substantiate this assessment.

Step 9: Prepare and submit the analysis report to support the MOS recruit accession quality requirement. An example analysis report is located in Appendix H. All DQ analysis reports must contain the following information in this format:

- a. Summary
 - (1) A statement of the Proponent's position on what the required recruit accession quality mix is for the MOS
 - (2) A concise enumeration of the reasons for this position
- b. Target Skill Level
 - (1) Identification of the Target Skill Level analyzed
 - (2) Discussion of the criteria used in selecting the Target Skill Level (i.e., reasons for selection)
- c. Performance measure(s)
 - (1) Narrative description of the performance measure(s)
 - (2) Detailed discussion of the validity and the reliability of the performance measure(s)
 - (3) Detailed justification of the minimum acceptable performance score
- d. Research design and implementation
 - (1) Explanation of the test administration and data collection procedures used
 - (2) Discussion of the statistical analysis and major findings
- e. Conclusion and recommendations
 - (1) The required recruit accession quality mix derived from following Steps 1-7
 - (2) An assessment of the probable impact on soldier quality requirements of projected changes in equipment, doctrine, training, and organization during the next five years.
 - (3) An estimate of how much the required recruit accession quality mix derived from following Steps 1-7 should be augmented or diminished based on these projected changes.

Figure 2 is a model depicting the procedures described above that Proponents must use to justify their accession quality requirements. The first eight steps (Step 9 being merely report preparation and submission) have been divided into three categories: empirical procedures (Steps 1-6), algorithmic procedures (Step 7), and estimation procedures (Step 8). Proponent efforts and expertise are essential to accomplishing the empirical procedures. It is in this category that Army accession quality justifications have usually fallen short. If the empirical procedures are not implemented correctly, the results of the algorithmic and the estimation procedures can be neither valid nor credible. On the other hand, there are a number of defensible ways to do the algorithmic procedures after the empirical procedures have been done. Due to the work of the

school combat developers, Proponents probably are in the best position to make the subjective assessments needed in the estimation procedures. In other words, the Proponents' highest DQ Program priority should be that of implementing the empirical procedures correctly.

PROPONENT QUALITY JUSTIFICATION MODEL

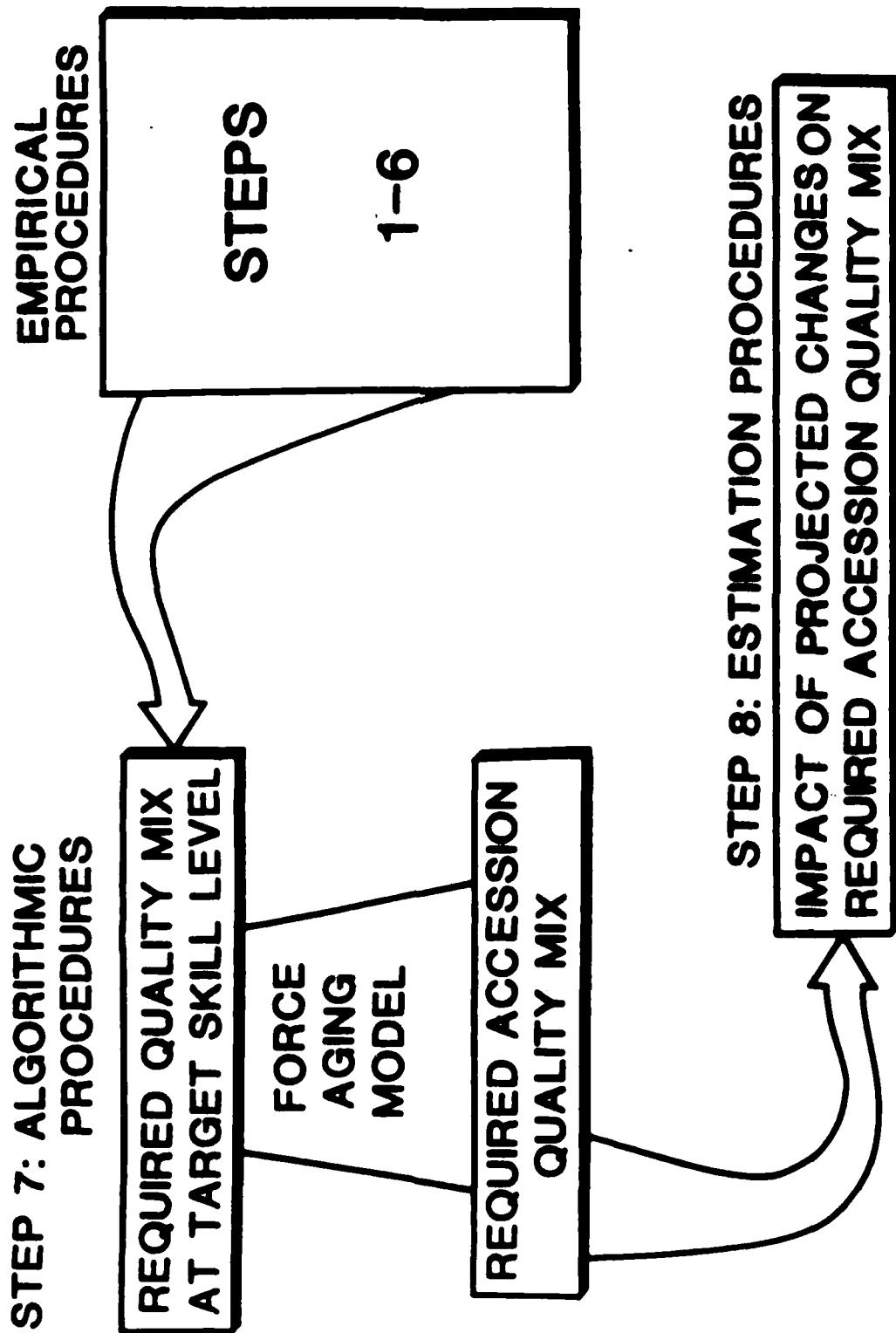


FIGURE 2

CHAPTER 3

DQ PROCESS AND MILESTONES

Collecting soldier quality and performance data in accordance with the specifications in this Handbook is a continuous process. Proponents should include in this process the ability to extract the data currently on hand to accomplish Steps 6-9 of Chapter 2 whenever a need arises. It is conceivable that unprogrammed data requests could cause a need for such information on any MDS at any time. Steps 1-4 should be reviewed and updated whenever changes in organization, personnel management policies, doctrine, training, equipment, etc., cause a major revision in an MDS. Implementation of Steps 1-5 should begin as soon as feasible after a new MDS is created.

USAREC usually works on two recruiting missions simultaneously: one for the current fiscal year and the other one for the next fiscal year under the Delayed Entry Program (DEP). USAREC prefers to recruit about half of next fiscal year's mission during the current fiscal year, if feasible. To have the maximum amount of influence on next year's mission (the second recruiting mission), TRADOC must present the fully justified Proponent recruit accession quality requirements to ODCSPER and negotiate the final DQ before the current fiscal year begins. To meet this objective, it is essential that the following milestones be met:

>May--Proponent DQ requirements for the second recruiting mission submitted to SEC-NCR

>July--TRADOC consolidated DQ requirements for the second recruiting mission submitted to ODCSPER

>September--ODCSPER, USAREC, and TRADOC agree on the most feasible DQ for the second recruiting mission

>October--USAREC starts recruiting the second half of the current fiscal year mission and the first half of the next fiscal year mission under the DEP

To clarify the milestone explanation above, the DQ for a specific fiscal year (FY91) is used below as an example:

>May 1989--Proponent DQ requirements for FY91 should be submitted to SEC-NCR

>July 1989--TRADOC consolidated DQ requirements for FY91 should be submitted to ODCSPER

>September 1989--ODCSPER, USAREC, and TRADOC should agree on the most feasible DQ for FY91

>October 1989--USAREC starts recruiting the second half of the FY90 mission and the first half of the FY91 mission under the DEP (Note: Agreement on the most feasible DQ for FY90 should have been reached in September 1988)

APPENDIX A

ARMED FORCES QUALIFICATION TEST (AFQT) CATEGORY TABLE

AFQT Category	Percentile Scores
I	93-100
II	65-92
IIIA	50-64
IIIB	31-49
IV	10-30
V	1-9

APPENDIX B

DETERMINING CONTENT VALIDITY

Background and Preview

Evaluating the validity of any performance measure is a question of the representativeness and adequacy of the criteria. Content validity is concerned with whether or not the performance measure contains a fair sample of the critical tasks. Since the procedure involves making inferences from a sample to a population, an evaluation of validity is made in terms of the adequacy of the performance measure. Operationally, content validity is the extent to which overlap exists between performance on the measure and ability to function in the defined job performance domain.

Step 1: Establish a content evaluation panel.

The content evaluation panel is composed of job incumbents, supervisors, trainers, and command staff.

Step 2: Identify critical tasks to be evaluated.

The content evaluation panel will initially review all existing materials relating to the MOS and define the critical tasks as well as the associated knowledge, skills, and abilities. The resultant list is distributed to each panel member for review.

Step 3: Evaluate test items/questions.

The resultant list of critical tasks is presented to each panel member with a set of test questions and independently indicates whether the knowledge, skill, or ability measured by each test question is:

- Essential;
- Useful but not essential; or
- Not necessary

to the performance of the critical task.

An item is rated essential if the soldier must possess this knowledge, skill, or ability to successfully perform the task. On the other hand, an item is rated useful if possessing this knowledge, skill, or ability would enhance the soldier's performance. Additionally, items rated as useful normally can be learned on the job. Items rated not necessary to the performance

of the task is self-explanatory. The format shown below may be useful in evaluating the items for relevancy.

Evaluation Sheet

(Name of Panel Member)

(Date)

(Performance Measure Title)

Please indicate by a check mark () your rating in the appropriate box.

Item Number	(Essential)	(Useful)	(Not necessary)
#1			
#2			
#3			

Step 4: Compute the content validity ratio.

Responses from all panelists are pooled and the number indicating "essential" for each item is determined. When all panelists indicate that an item is essential or not essential, one can have confidence that the item is or is not truly essential. It is when there is no consensus that problems arise. Therefore, the following assumption is made: any item which is perceived to be essential by more than half (beyond 50%) of the panelists has some degree of validity and will be counted as essential. A content validity ratio is then computed for each test item. The formula for computing the CVR is:

$$= \frac{ne - (N/2)}{(N/2)}$$

where "ne" is the number of panelists indicating essential and N is the total number of panelists.

Step 5: Test each item for statistical significance.

Test items are eliminated if the CVR fails to meet statistical significance. For example, when the content evaluation panel is composed of 15 members, a minimum CVR of .49 is required to

satisfy the .05 level of significance. Only those items with CVR values meeting/exceeding this minimum are retained in the test.

Minimum values of CVR
(One-tailed test, $p = .05$)

Total number of panelists	Minimum value
5	.99
6	.99
7	.99
8	.75
9	.78
10	.62
11	.59
12	.56
13	.54
14	.51
15	.49
20	.42
25	.37
30	.33
35	.31
40	.29

(This table is taken from "Personnel Psychology" (Vol 28, 563-575). C. H. Lawshe.)

Step 6: Compute the content validity index.

The CVI is computed for the whole test. The CVI is the mean of the CVR values of the retained items meeting the minimum CVR tabled value. The formula for computing the CVI is:

$$= \frac{(\text{sum CVR})}{N}$$

APPENDIX C

DETERMINING RELIABILITY

Background and Preview

Reliability estimates indicate the degree of consistency of the performance measure. The technique used in this methodology, the Kuder-Richardson reliability coefficient, involves an analysis of item variance which indicates the degree to which the various test items on a test are intercorrelated.

The formula for computing the reliability coefficient is:

$$r_{KR} = \frac{n}{n-1} \left[\frac{sx^2 - (\sum pq)}{sx^2} \right]$$

where n is the number of items on the test, and sx^2 is the variance of the total scores on the test. The final term ($\sum pq$) is found by computing the proportion of the group who pass (p) and do not pass (q) each item, where $q=1-p$. The product of p and q is then computed for each item, and these products are added for all items to yield $\sum pq$.

The formula for computing the variance is:

$$sx^2 = \frac{\sum (x^2) - \frac{(\sum x)^2}{n}}{n}$$

Step 1: Compute an index of internal consistency.

Suppose that a ten-item test was administered to ten people. Each item was scored either 1 or 0; 1 for a correct response, 0 for an incorrect response. The results were:

People	Item										X
	1	2	3	4	5	6	7	8	9	10	
A	1	1	1	1	1	1	1	1	1	1	10
B	1	1	1	1	1	1	1	1	1	0	9
C	1	1	1	1	1	1	1	0	1	0	8
D	1	1	1	1	1	1	1	1	0	0	8
E	1	1	1	1	1	0	1	0	0	1	7
F	1	1	1	1	1	1	0	1	0	0	7
G	1	1	1	1	1	0	0	1	0	0	6
H	1	1	1	1	1	0	0	0	1	0	6
I	1	1	1	1	0	1	0	0	0	0	5
J	1	1	1	0	0	1	0	0	0	0	4
sum	10	10	10	9	8	7	5	5	4	2	70

The proportion passing each item (p) is the number of people passing the item (n) divided by the total number in the sample (n=10). The proportion failing the item (q) equals 1-p. These data are shown below:

item	1	2	3	4	5	6	7	8	9	10
	10	10	10	9	8	7	5	5	4	2
p	1.0	1.0	1.0	0.9	0.8	0.7	0.5	0.5	0.4	0.2
q	0.0	0.0	0.0	0.1	0.2	0.3	0.5	0.5	0.6	0.8
pq	.00	.00	.00	.09	.16	.21	.25	.25	.24	.16

The figure needed for the formula (sum pq) is the product of pq summed over all ten items:

$$\text{sum pq} = .09 + .16 + .21 + .25 + .25 + .24 + .16 = 1.36$$

Step 2: Calculate sx^2 .

To calculate sx, write down in column form each case and the corresponding score.

	1	2
Cases	x^2	x
A		10
B		9
C		8
D		8
E		7
F		7
G		6
H		6
I		5
J		4
n=	sum (x^2)	(sum x)
		(sum x) ²

Step 3: Fill in the table.

a. Fill in column 1 by squaring the x score in column 2. At the bottom of column 1, add up all the values (this is sum (x^2) from the formula).

b. Complete column 2 by adding up all the x scores. This is (sum x) from the formula. Underneath write down (sum x)².

c. A completed table is shown below.

	1	2
Cases	x^2	x
A	100	10
B	81	9
C	64	8
D	64	8
E	49	7
F	49	7
G	36	6
H	36	6
I	25	5
J	16	4
	520	70
		4900

Compute sx^2

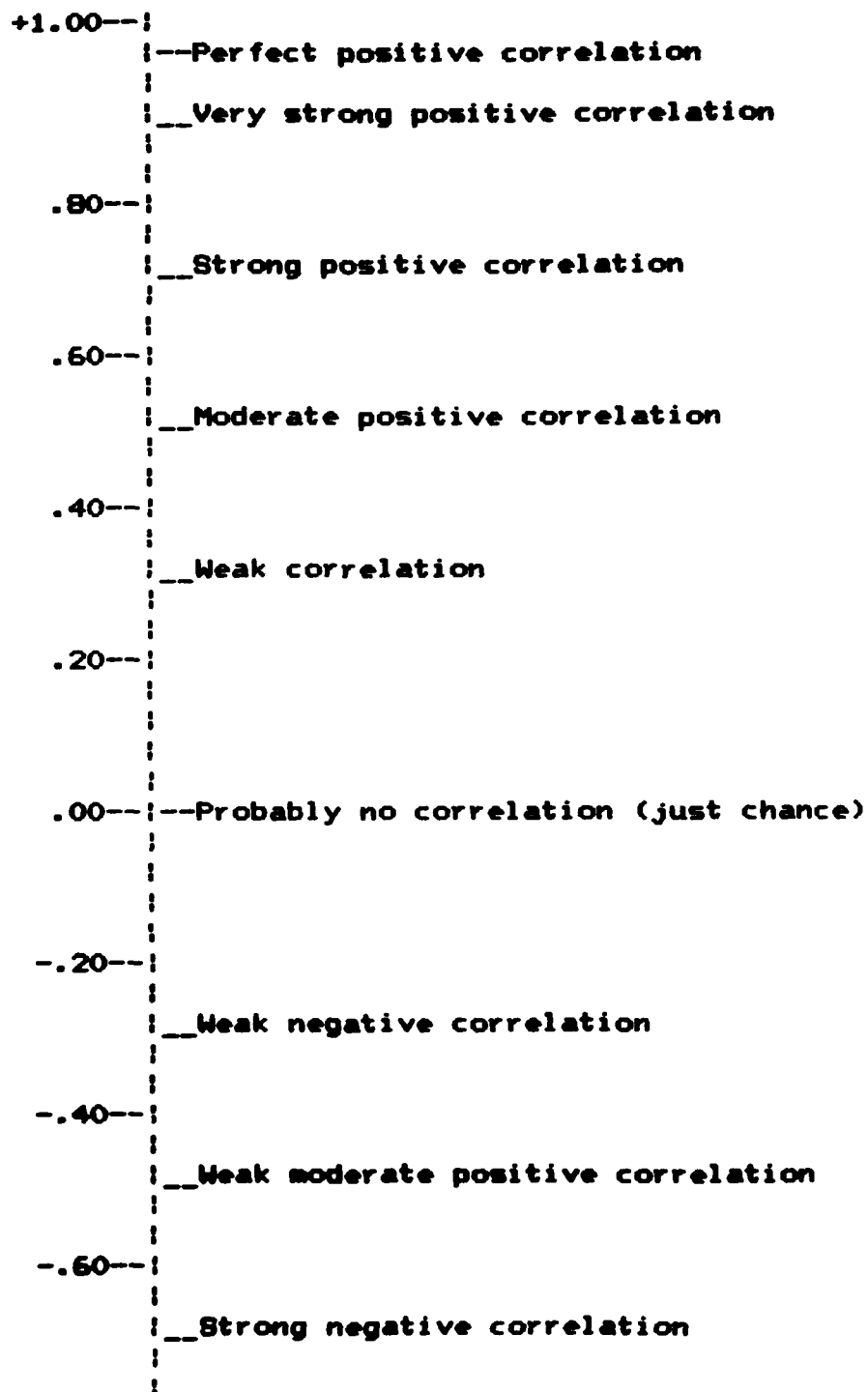
$$\begin{aligned}
 &= \frac{\sum(x^2) - (\sum x)^2}{n} \\
 &= \frac{520 - \frac{4900}{10}}{10} \\
 &= \frac{(520 - 490)}{10} \\
 &= 3.00
 \end{aligned}$$

Substituting in the formula

$$\begin{aligned}
 r_s &= \frac{n}{n-1} \left[\frac{sx^2 - (\sum x)^2}{sx^2} \right] \\
 &= \frac{10}{10-1} \left[\frac{3.00 - 1.25}{3.00} \right] \\
 &= \frac{10}{9} \left[\frac{1.75}{3.00} \right] \\
 &= .6074
 \end{aligned}$$

Interpret the statistic.

The table below provides the range of possible correlations and their usual interpretations.



-.80--|
|
|__Very strong negative correlation
|
-1.00--|---Perfect negative correlation

APPENDIX D

TESTING THE SIGNIFICANCE OF r

This section discusses how much confidence to place in a particular correlation. The recommended approach is to establish confidence limits around the coefficient. These limits indicate the range of probable correlations on subsequent runs of the performance test.

The significance of an obtained r may be tested against the hypothesis that the population r is in fact zero. If the computed r is large enough to invalidate or cast serious doubt upon the null hypothesis, accept r as indicating the presence of at least some degree of correlation.

To make the test, enter the table below with $(N-2)$ degrees of freedom and compare the obtained r with the tabulated entries. For example, when $N=10$, an r must be .632 and .765 to be significant at the .05 and .01 levels of significance, respectively. This means that only 5 times in 100 trials would an r as large as +.632 arise from fluctuations of sampling alone if the population r were actually .00 and only once in 100 trials would an r as large as +.765 appear if the population r were .00. It is clear then that the obtained r of .6074, since it is smaller than the tabled values, is not significant at the .05 or .01 levels of significance.

Correlation coefficients at the .05/.01 levels of significance

Degrees of Freedom (N-2) .05 .01			Degrees of Freedom (N-2) .05 .01			Degrees of Freedom (N-2) .05 .01		
1	.997	1.000	16	.468	.590	35	.325	.418
2	.950	.990	17	.456	.575	40	.304	.393
3	.878	.959	18	.444	.561	45	.288	.372
4	.811	.917	19	.433	.549	50	.273	.354
5	.754	.874	20	.423	.537	60	.250	.325
6	.707	.834	21	.413	.526	70	.232	.302
7	.666	.798	22	.404	.515	80	.217	.283
8	.632	.765	23	.396	.505	90	.205	.267
9	.602	.735	24	.388	.496	100	.195	.254
10	.576	.708	25	.381	.487	125	.174	.228
11	.553	.684	26	.374	.478	150	.159	.208
12	.532	.661	27	.367	.470	200	.138	.181
13	.514	.641	28	.361	.463	300	.113	.148
14	.497	.623	29	.355	.456	400	.098	.128
15	.482	.606	30	.349	.449	500	.088	.081

(This table is taken from Garrett, H. E., "Statistics in Psychology and Education," (6th ed.; New York: David McKay, 1966).

APPENDIX E

DETERMINING THE MINIMUM ACCEPTABLE SCORE

Background and Preview

When a continuous test is used to distinguish among members of a criterion group, some test score must be chosen as the cutting point. Selecting a cutting point is often difficult and is typically done arbitrarily based on some standard that is seen as reasonable. Ideally, the cut score should be flexible, not absolute, and should consider the test score distributions of the criterion sample.

By establishing and quantifying the relationship and dependencies between variables, a cut-off score based on various levels of ASVAB performance can be estimated. One way of estimating the cut score is to fit a straight line to the observations.

The general form for the equation of a straight line is:

$$y = a + bx$$

where "y" is defined as the variable whose value needs to be determined and "x" represents the various levels of ASVAB performance.

$b = r(s_y/S)$ b represents the slope of the line.

$a = \frac{(\sum y)}{n} - b \frac{(\sum x)}{n}$ a represents the y-intercept (a constant value).

$s_x = \sqrt{\frac{\sum(x^2) - \frac{(\sum x)^2}{n}}{n}}$ standard deviation of the quality data (AFQT or AA score).

$s_y = \sqrt{\frac{\sum(y^2) - \frac{(\sum y)^2}{n}}{n}}$ std dev of the performance measure.

$r = \frac{[n(\sum xy)] - [(\sum x)(\sum y)]}{\sqrt{[n \cdot \sum(x^2) - (\sum x)^2][n \cdot \sum(y^2) - (\sum y)^2]}}$ correlation

$r_u = \frac{r(S/s_x)}{\sqrt{(1 - r^2) + [r^2(S^2/s_x^2)]}}$ corrected correlation coefficient for range restriction (in the predictor variable-- AA scores)

where S is the standard deviation of the population and s_x is the standard deviation of the sample. (In this methodology S is

where S is the standard deviation of the population and s_x is the standard deviation of the sample. (In this methodology S is a constant value which is equal to 20 since Area aptitude scores have a mean of 100 and a standard deviation of 20.)

Step 1: Prepare a table.

To calculate the straight line formula, start with two sets of scores from one set of cases. Write down in column form the performance measure score and the corresponding AA or AFBT score for each individual.

	1	2	3	4	5
Cases	x^2	x	xy	y	y^2
		102		79	
		116		85	
		97		81	
		106		94	
		96		89	
		100		86	
		106		87	
		84		77	
		117		95	
		120		92	
		90		90	
		106		86	
		112		84	
		112		86	
		119		90	
		113		81	
n	$\text{sum}(x^2)$	$\text{sum } x$ $(\text{sum } x)^2$	$\text{sum}(xy)$	$\text{sum } y$ $(\text{sum } y)^2$	$\text{sum}(y^2)$

Step 2: Fill in the table.

- Fill in column 1 by squaring the x-score in column 2. At the bottom of column 1 add up the values. This is $\text{sum}(x^2)$.
- Complete column 2 by adding up all the x-scores. This is $\text{sum } x$. Compute $(\text{sum } x)^2$.
- Fill in column 3 by multiplying the x-score in column 2 by the y-value in column 4. At the bottom of column 3 add up all the values. This is $\text{sum } xy$.
- Complete column 4 by adding up all the y-values. This is $\text{sum } y$. Compute $(\text{sum } y)^2$.

e. Fill in column 5 by squaring the y-values in column 4. At the bottom of column 5, add up all the values. This is $\sum y^2$.

(A completed table is shown on the following page).

	1	2	3	4	5
Cases	x^2	x	xy	y	y^2
1	10404	102	8088	79	6241
2	13456	116	9860	85	7225
3	9409	97	7857	81	6561
4	11236	106	9964	94	8836
5	9216	96	8544	89	7921
6	10000	100	8600	86	7396
7	11236	106	9222	87	7569
8	7056	84	6468	77	5929
9	13689	117	11115	95	9025
10	14400	120	11040	92	8464
11	8100	90	8100	90	8100
12	11236	106	9116	86	7396
13	12544	112	9408	84	7056
14	12544	112	9632	86	7396
15	14161	119	10710	90	8100
16	12769	113	9153	81	6561
n=16	181456	1696	146847	1382	119776
		22876416		1909924	

Step 3: Compute r.

$$\begin{aligned}
 &= \frac{[n(\sum xy)] - [(\sum x)(\sum y)]}{\sqrt{[n \sum(x^2) - (\sum x)^2][n \sum(y^2) - (\sum y)^2]}} \\
 &= \frac{16(146847) - (1696)(1382)}{\sqrt{[16(181456) - 2876416][16(119776) - 1909924]}} \\
 &= \frac{2349552 - 2343872}{\sqrt{(2903296 - 2876416)(1916416 - 1909924)}} \\
 &= \frac{5680}{\sqrt{(26880)(6492)}} \\
 &= \frac{5680}{\sqrt{174504960}} \\
 &= \frac{5680}{13210.03} \\
 &= .4300
 \end{aligned}$$

Step 4: Compute s_x and s_y .

$$\begin{aligned}
 s_x &= \sqrt{\frac{\sum(x^2) - (\sum x)^2}{n}} \\
 &= \sqrt{\frac{181486 - \frac{2875416}{16}}{16}} \\
 &= \sqrt{\frac{181486 - 179776}{16}} \\
 &= \sqrt{\frac{1710}{16}} \\
 &= \sqrt{105} \\
 &= 10.2470
 \end{aligned}$$

$$\begin{aligned}
 s_y &= \sqrt{\frac{\sum(y^2) - (\sum y)^2}{n}} \\
 &= \sqrt{\frac{119776 - \frac{1209924}{16}}{16}} \\
 &= \sqrt{\frac{119776 - 119370.25}{16}} \\
 &= \sqrt{\frac{405.75}{16}} \\
 &= \sqrt{25.3594} \\
 &= 5.0358
 \end{aligned}$$

Step 5: Correct r for range restriction.

$$\begin{aligned}
 r_u &= \frac{r(S/s_x)}{\sqrt{(1-r^2) + [r^2(S^2/s_x^2)]}} \\
 &= \frac{.43(20/10.247)}{\sqrt{[1 - (.43)^2] + (.43)^2[(20)^2/(10.247)^2]}} \\
 &= \frac{(.43)(1.9518)}{\sqrt{[1 - .1849] + [(.1849)(400/105)]}} \\
 &= \frac{.8393}{\sqrt{.8151 + .7044}} \\
 &= \frac{.8393}{\sqrt{1.5195}} \\
 &= \frac{.8393}{1.2327} \\
 &= .6809
 \end{aligned}$$

Step 6: Determine the significance of the corrected r .

Mathematically, a more defensible method of testing the significance of an r , especially when the coefficient is very high or very low, is to convert r into a z -function and find the standard error of z . The z -function has the advantage over r in that its standard error depends only upon the size of the sample (N) and is independent of the size of r .

The formula for the standard error of z is:

$$SE = \frac{1}{\sqrt{N-3}}$$

Suppose that $r = .6809$ and $N = 16$. According to the table shown below, an r of .6809 corresponds to a z of .83.

Conversion of an r into a corresponding z -coefficient

r	z	r	z	r	z
.25	.26	.60	.69	.90	1.47
.26	.27	.61	.71	.905	1.50
.27	.28	.62	.73	.910	1.53
.28	.29	.63	.74	.915	1.56
.29	.30	.64	.76	.920	1.59
.30	.31	.65	.78	.925	1.62
.31	.32	.66	.79	.930	1.66
.32	.33	.67	.81	.935	1.70
.33	.34	.68	.83	.940	1.74
.34	.35	.69	.85	.945	1.78
.35	.37	.70	.87	.950	1.83
.36	.38	.71	.89	.955	1.89
.37	.39	.72	.91	.960	1.95
.38	.40	.73	.93	.965	2.01
.39	.41	.74	.95	.970	2.09
.40	.42	.75	.97	.975	2.18
.41	.44	.76	1.00	.980	2.30
.42	.45	.77	1.02	.985	2.44
.43	.46	.78	1.05	.990	2.65
.44	.47	.79	1.07	.995	2.99
.45	.48	.80	1.10		
.46	.50	.81	1.13		
.47	.51	.82	1.16		
.48	.52	.83	1.19		
.49	.54	.84	1.22		
.50	.55	.85	1.26		
.51	.56	.86	1.29		
.52	.58	.87	1.33		
.53	.59	.88	1.38		
.54	.60	.89	1.42		
.55	.62				
.56	.63				
.57	.65				
.58	.66				
.59	.68				

r 's under .25 may be taken as equivalent to z 's.

(This table is taken from Fisher, R. A., "Statistical Methods for Research Workers", (8th ed.; London: Oliver and Boyd, 1941).

Substituting in the formula for SE :

$$\begin{aligned}
 &= \frac{1}{\sqrt{N-3}} \\
 &= \frac{1}{\sqrt{16-3}} \\
 &= .2774
 \end{aligned}$$

The 95% confidence interval for the true ρ is computed using the formula [$\bar{r}$ (SE)], where z is the z -score for the .95 level of confidence. The confidence interval for the true ρ is 0.2864 to 1.3736 (i.e., $.83 \pm 1.96 \times .2774$ or $.83 \pm .5436$).

Converting these z 's back into r 's yields a confidence interval from .27 to .88.

Interpret the statistic.

The 95% confidence interval for ρ extends from the .27 to .88. Thus, there is a 95% likelihood that this interval includes the true ρ . Since the interval includes only positive values and excludes zero, the correlation is considered to be statistically significant.

Step 7: Compute b.

$$\begin{aligned}
 b &= r(s_y/s) \\
 &= .6809(5.0358/20) \\
 &= .6809(.2518) \\
 &= .1714
 \end{aligned}$$

Step 8: Compute a.

$$\begin{aligned}
 a &= \frac{(\sum y)}{n} - b \frac{(\sum x)}{n} \\
 &= (1382/16) - (.1714) \left[\frac{1696}{16} \right] \\
 &= 86.3750 - (.1714)(106) \\
 &= 86.3750 - 18.1730 \\
 &= 68.2020
 \end{aligned}$$

Select the cut point.

a. Selecting the cut score is based on the established ASVAS standard (selector area aptitude score) for entry into the NDS. Suppose that the selector AA score for entry into the tested NDS is 100. Substituting values in the following formula yields:

$$\begin{aligned} y &= a + bx = 68.2020 + .1714x \\ &= 68.2020 + .1714(100) \\ &= 85.3 \text{ (predicted cut score)} \end{aligned}$$

b. When selecting the cut score, analysts should examine the impact of various levels of predicted cut scores.

(1) Construct a frequency chart with values of the performance measure. Each cell gives the frequency of data falling at the performance measure ranges.

Frequency Chart

Performance Scores	I-IIIA	IIIB	Overall
65-69	1	0	1
70-74	1	0	1
75-79	1	2	3
80-84	3	5	8
85-89	4	11	15
90-94	8	4	12
95-99	19	6	25
	37	28	65

Cut score = 85

(2) Convert to conditional probabilities

It can be seen from the above table that if the cut score is set at 85, an individual in AFBT Categories I-IIIA can be expected to have an 84% (31/37) chance of being successful; a IIIB can be expected to have a 75% (21/28) chance of being successful. A completed probability table is shown below.

Cut Score	I-IIIA	IIIB	Total
65	1.00	1.00	1.00
70	.97	1.00	.98
75	.95	1.00	.97
80	.92	.93	.92
85	.84	.75	.80
90	.73	.36	.57
95	.51	.21	.38

APPENDIX F

TESTING THE STATISTICAL SIGNIFICANCE OF THE DIFFERENCE BETWEEN TWO MEANS

Background and Preview

Testing is the part of the scientific method that provides the basis for accepting or rejecting a hypothesis. The test criteria allow formulation of a hypothesis (soldiers with higher AFQT scores perform better than those with lower scores). The following steps explain how to determine if the difference between the groups is statistically significant. The formula for calculating the t-statistic is:

$$t = \frac{a-b}{sd}$$

where a and b represent the mean scores of the two groups on the performance measure and sd is the standard deviation.

Step 1: Prepare the input data.

For each group, compute the mean score and standard deviation and record these values in the descriptive table below along with the number, n, of scores of each group.

Group	Mean Score	Std Dev	n
a (1-3A)	74.7	6.84	25
b (3B)	72.4	8.33	25

Step 2: Calculate the difference between group means, d.

$$\begin{aligned}d &= a-b \\&= 74.7-72.4 \\&= 2.3\end{aligned}$$

Step 3: Calculate the standard deviation, sd.

$$\begin{aligned}sd &= \sqrt{\frac{sa (na-1)+sb (nb-1)}{na+nb-2} \left[\frac{1}{na} + \frac{1}{nb} \right]} \\&= \sqrt{\frac{(6.84)(25-1)+(8.33)(25-1)}{25+25-2} \left[(1/25 + 1/25) \right]}\end{aligned}$$

$$\begin{aligned}
&= \sqrt{\left[\frac{(6.84)(25-1) + (8.33)(25-1)}{48} \right] (2/25)} \\
&= \sqrt{\left[\frac{(46.7856)(24) + (69.3889)(24)}{48} \right] (2/25)} \\
&= \sqrt{\frac{(1122.8544 + 1665.3336)(.08)}{48}} \\
&= \sqrt{(58.0873)(.08)} \\
&= \sqrt{4.647} = 2.1557
\end{aligned}$$

Step 4: Calculate t.

$$= 2.3/2.1557$$

$$= 1.0669 \text{ (obtained t-value)}$$

Step 5: Find the tabled t-statistic.

- Find in column 1 the value calculated for $na+nb-2$.
- Find the t-value which corresponds to column 1.

Tabled t-Values
(p = .05)

Column 1	Column 2	Column 1	Column 2
6	2.45	21	2.08
7	2.36	22	2.07
8	2.31	23	2.07
9	2.26	24	2.06
10	2.23	25	2.06
11	2.20	26	2.06
12	2.18	27	2.05
13	2.16	28	2.05
14	2.14	29	2.04
15	2.13	30	2.04
16	2.12	40	2.02
17	2.11	60	2.00
18	2.10	120	1.98
19	2.09		1.96
20	2.09		

(This table is taken from Table III of Fisher and Yates: "Statistical Tables for Biological, Agricultural, and Medical Research," published by Longman Group, Ltd).

c. $t = 2.02$ (tabled t-statistic)

Step 6: Interpret the statistic.

The obtained t-value (step 4) may have been negative. Ignore the negative sign if there is one, and compare the absolute size of the obtained value. If the obtained value is larger than the tabled value (step 5), the difference between the two groups was statistically significant.

If the obtained t-value was not larger than the tabled t-value, the difference could have been the result of chance. Since the result cannot be regarded as showing that either group did better than the other, the difference between the two groups was not statistically significant.

Step 7: Calculate the 95% confidence limits.

a. Figure the lower limit.

$$LL = d - (t)(sd)$$

where d is the value calculated in step 2; sd calculated in step 3; and the tabled t-value from step 5. Substituting the values from the previous steps yield:

$$LL = 2.3 - (2.02)(2.1557)$$

$$= -2.0545$$

b. Figure the upper limit.

$$UL = d + (t)(sd)$$

$$= 2.3 + (2.02)(2.1557)$$

$$= 6.6545$$

Step 8: Interpret the statistic.

Does the confidence interval (that is, the range of values from the lower to the upper limit) include zero? If YES, the obtained difference between the groups could have been due to chance. The difference between the groups is not statistically significant.

If NO and both limits are either positive or negative, the difference between means is likely to occur again in favor of the same group if the program were repeated. The obtained difference is statistically significant. In other words, it is extremely unlikely that this difference occurred by chance.

The AFQT group I-IIIA NCO's generally achieved higher scores than the AFQT group IIIB NCO's. The obtained difference was not statistically significant. However, this could have been

obtained by chance (since the range of values includes zero). There is a high probability (more than 95%) that the difference of 2.3 points for the I-IIIA NCO's could have resulted from chance variation between the groups.

APPENDIX G

ESTABLISHING DQ REQUIREMENTS

Background and Preview

Defining manpower quality in measurable terms is critical to the Army's ability to estimate accession requirements. Estimating entry level quality needs in terms of skill level 1 or 2 is not sufficient because the Army must develop its supervisors from its pool of recruits. A sufficient number of individuals must be accessed, trained, and retained from the recruit pool to fill future NCO needs.

This appendix contains a step-by-step explanation for translating the performance criteria for a given skill level to an area aptitude distribution and then to an AFQT distribution that would be required at lower skill levels.

A very basic assumption underlying this approach is that the sample of soldiers administered the performance measure is representative of the population. A representative sample is defined as one in which the distribution of ASVAB scores closely parallels that of the population.

Additionally, a sufficient number of soldiers must be tested to permit statistically valid findings. In other words, a sample is tested and the results are projected to the population.

Step 1: Construct a frequency table.

Using the selector area aptitude scores obtained from the sample, construct a frequency table with values of the selector area aptitude. Each cell of the frequency table gives the number of data cases falling at each area aptitude ranges.

Selector area aptitude composite (GM)

90-94	95-99	100-104	105-109	110-114	115-119	120-124
9	8	13	17	9	7	2

Step 2: Convert the AA distribution into an AFQT distribution.

Convert the AA distribution into an AFQT distribution using the attached conversion tables, one for each AA composite.

To use the tables, multiply the sample size of each AA score range by the percentages in the appropriate column for the AA

score range. This process results in the expected AFQT distribution for the individuals in the sample.

For example, using the AA distribution identified above and the AA to AFQT distribution table for the GM composite, the following AFQT distribution is obtained:

	90- 94	95- 99	100- 104	105- 109	110- 114	115- 119	120- 124	%
Expected AFQT: distribution :	9	8	13	17	9	7	2	
I-IIIA	3	4	8	12	7	6	2	65
IIIB	3	2	4	4	2	1	0	25
IV	3	2	1	1	0	0	0	10

Notes:

1. The numbers have been rounded off to the nearest whole number.
2. If 9 individuals scored between 90-94 on the GM composite, 2.81 individuals (9x.3125) were projected to be in AFQT category I-IIIA, 3.47 (9x.3855) to be in AFQT category IIIB, and 2.72 (9x.3019) in AFQT category IV.

Step 3: Determine the accession quality mix.

Translate the obtained AFQT distribution (step 2) into the AFQT distribution that would be required for accession.

- a. Based on the performance data, the following skill level 3 AFQT distribution was obtained:

Per cent by AFQT category*

I-IIIA	IIIB	IV
65	25	10

***desired quality mix**

- b. To translate these numbers into the accession vector required to yield the desired quality mix, divide each percentage by the historical continuation/survival rates.

I-IIIA	65/.0645=	993.88
IIIB	25/.0696=	359.20
IV	10/.0837=	119.47

c. To determine the accession proportions, sum the numbers and then divide each number by the sum.

$$993.88/1472.55 = .67**$$

$$359.20/1472.55 = .24**$$

$$119.47/1472.55 = .09**$$

**desired accession quality mix

CONVERSION TABLES FOR ESTIMATING AFQT PERCENTILE DISTRIBUTIONS FROM AREA APTITUDE STANDARD SCORE DISTRIBUTIONS:

Clerical (CL) Aptitude Area Composite

AA Standard Score

AFQT %ile	80- 84	85- 90	90- 94	95- 99	100- 104
I-III A	.00	.10	6.11	27.64	62.78
IIIB	7.47	38.95	76.19	67.37	36.80
IV	92.53	60.94	17.70	5.00	.41

Clerical (CL) Aptitude Area Composite

AA Standard Score

AFQT %ile	105- 109	110- 115	115- 120	120- 124	125- 129	130>
I-III A	88.23	97.26	98.83	100.00	100.00	100.00
IIIB	11.77	2.73	1.17	.00	.00	.00
IV	.00	.00	.00	.00	.00	.00

Army Combat (CD) Aptitude Area Composite

AA Standard Score

AFQT %ile	80- 84	85- 89	90- 94	95- 99	100- 104
I-III A	2.31	12.63	24.29	43.49	62.48
IIIB	30.02	38.02	44.11	38.87	25.31
IV	67.67	49.36	31.58	17.63	12.20

Army Combat (CD) Aptitude Area Composite

AA Standard Score

AFQT %ile	105- 109	110- 115	115- 120	120- 124	125- 129	130>
I-III A	72.89	82.08	87.50	95.56	99.03	100.00
IIIB	21.16	14.57	11.76	4.44	.97	.00
IV	5.96	3.36	.74	.00	.00	.00

Electronics (EL) Aptitude Area Composite

AA Standard Score

AFQT %ile	80- 84	85- 89	90- 94	95- 99	100- 104
I-III A	1.52	6.12	21.89	45.42	58.68
IIIB	25.55	45.80	48.10	42.36	36.49
IV	72.93	48.09	30.01	12.22	4.84

Electronics (EL) Aptitude Area Composite

AA Standard Score

AFQT %ile	105- 109	110- 115	115- 120	120- 124	125- 129	130>
I-III A	77.84	92.46	96.13	98.72	99.95	100.00
IIIB	19.75	7.26	3.87	1.27	.06	.00
IV	2.40	.28	.00	.00	.00	.00

Field Artillery (FA) Aptitude Area Composite

AA Standard Score

AFQT %ile	80- 84	85- 89	90- 94	95- 99	100- 104
I-III A	1.22	6.13	20.75	38.76	60.50
IIIB	27.76	38.01	50.54	48.49	36.44
IV	71.03	55.84	28.70	12.75	3.06

Field Artillery (FA) Aptitude Area Composite**AA Standard Score**

AFQT	105-	110-	115-	120-	125-	130>
%ile	109	115	120	124	129	
I-III A	80.88	90.01	97.55	100.00	100.00	100.00
IIIB	17.62	9.59	2.45	.00	.00	.00
IV	1.50	.41	.00	.00	.00	.00

General Maintenance (GM) Aptitude Area Composite**AA Standard Score**

AFQT	80-	85-	90-	95-	100-
%ile	84	89	94	99	104
I-III A	6.53	18.09	31.25	50.94	58.78
IIIB	32.29	39.57	38.55	27.69	28.66
IV	61.16	42.34	30.19	21.38	12.56

General Maintenance (GM) Aptitude Area Composite**AA Standard Score**

AFQT	105-	110-	115-	120-	125-	130>
%ile	109	115	120	124	129	
I-III A	67.88	78.09	87.73	93.01	95.96	99.13
IIIB	23.00	17.34	11.25	5.68	4.04	.86
IV	9.14	4.56	1.01	1.32	.00	.00

Mechanical Maintenance (MM) Aptitude Area Composite**AA Standard Score**

AFQT	80-	85-	90-	95-	100-
%ile	84	89	94	99	104
I-III A	7.90	15.05	31.40	52.79	63.23
IIIB	25.19	42.17	39.80	28.09	23.21
IV	66.92	42.78	28.81	19.11	13.54

Mechanical Maintenance (PM) Aptitude Area Composite**AA Standard Score**

AFQT	105-	110-	115-	120-	125-	130>
Xile	109	115	120	124	129	
I-III A	69.25	75.35	79.06	89.32	94.11	98.58
IIIB	21.09	17.70	18.76	8.62	4.84	1.42
IV	9.67	6.96	2.19	2.07	1.05	.00

Operators/Food (OF) Aptitude Area Composite**AA Standard Score**

AFQT	80-	85-	90-	95-	100-
Xile	84	89	94	99	104
I-III A	.18	2.85	15.47	34.53	59.44
IIIB	11.44	36.11	50.01	44.04	32.14
IV	88.38	61.07	34.51	21.43	8.42

Operators/Food (OF) Aptitude Area Composite**AA Standard Score**

AFQT	105-	110-	115-	120-	125-	130>
Xile	109	115	120	124	129	
I-III A	73.58	80.71	89.95	96.08	99.87	100.00
IIIB	22.65	16.35	9.41	3.92	.14	.00
IV	3.76	2.93	.63	.00	.00	.00

Surveillance/Communications (SC) Aptitude Area Composite**AA Standard Score**

AFQT	80-	85-	90-	95-	100-
Xile	84	89	94	99	104
I-III A	1.16	6.41	22.55	45.72	66.18
IIIB	29.73	45.79	50.12	38.36	22.69
IV	69.11	47.81	27.35	15.91	11.11

Surveillance/Communications (SC) Aptitude Area Composite**AA Standard Score**

AFBT Xile	105- 109	110- 115	115- 120	120- 124	125- 129	130>
I-III A	71.08	77.77	87.82	95.32	99.99	99.91
IIIB	20.83	19.87	12.12	4.68	.00	.00
IV	8.09	2.36	.06	.00	.00	.00

Skilled Technical (ST) Aptitude Area Composite**AA Standard Score**

AFBT Xile	80- 84	85- 89	90- 94	95- 99	100- 104
I-III A	1.05	4.52	18.34	36.01	60.26
IIIB	22.16	35.67	51.59	48.33	34.30
IV	76.80	59.81	30.07	15.66	5.45

Skilled Technical (ST) Aptitude Area Composite**AA Standard Score**

AFBT Xile	105- 109	110- 115	115- 120	120- 124	125- 129	130>
I-III A	73.03	87.56	96.41	98.39	99.81	99.33
IIIB	24.81	11.24	3.59	1.61	.19	.66
IV	2.17	1.19	.00	.00	.00	.00

APPENDIX H

SAMPLE DB ANALYSIS REPORT

I. Target Skill Level: E6s (skill level 3) have the required technical expertise and are usually first line supervisors or section chiefs.

II. Performance Measure

a. Basic NCO course (BNCOC)

(1) Length: 12 weeks

(2) Common subjects: Leadership, training management, nuclear, chemical, biological, maintenance, land navigation, logistics, personnel, and weapons.

(3) Narrative description: Basic level training is functional within the career management field (CMF) and is oriented toward appropriate qualification at section chief skill level. Emphasis is placed on leadership and human relations skills and knowledge of subjects required to perform effectively at the section chief level. Training is directed toward providing a firm comprehension of NCO's role in combat, combat support, and service support.

b. Content validity/reliability

(1) The Content Evaluation Panel consisted of 40 job incumbents, supervisors, and instructors. Each panel member was supplied with a copy of the end-of-course comprehensive examination and a copy of the critical task list along with the required knowledge, skills, and abilities associated with each critical task. The "essentiality" question was modified by changing the second response "useful but not essential," to "useful but can be learned on-the-job." Each panelist indicated his response on the answer sheet for each of the 100 questions.

(2) Quantifying the results. The responses were pooled for each item and the content validity ratio was computed for each (These CVR values are tabulated below). When the mean of the values was computed, the content validity index for the total test is .53; when the items that failed to meet statistical significance are eliminated the CVI is .74.

Distribution of CVR Values

no	CVR			cumulative
40	1.00	15	15.00	15.00
37	.85	14	11.90	26.90
34	.70	13	9.10	36.00
32	.60	12	7.20	43.20
29	.45	11	4.95	48.15
25	.25	10*	2.50	50.65
24	.20	9*	1.80	52.45
21	.05	8*	.40	52.85
20	.00	7*	.00	52.85

$$CVI = 52.85/100 = .5285$$

* Failed to meet statistical significance--a minimum CVR of .29 is required to satisfy the .05 level of significance.

(3) Kuder-Richardson reliability coefficient:

$$r_k = \frac{(100/99)(49.3931 - 9.0598)}{49.3931}$$

$$= .8248 \text{ (strong positive correlation)}$$

Test for significance

When N=65 and df is 63, an r must be .250 to be significant at the .05 level and .325 to be significant at the .01 level. It is clear that the obtained r of .8248 is highly significant at the .01 level, since it is higher than .325.

c. Minimum acceptable score

Correlation coefficient (AA score vs EOC exam)

$$r = \frac{65(618524) - (6836 \times 5872)}{\sqrt{[65(724230) - (6836)^2][65(533678) - (5872)^2]}}$$

$$= .2354$$

Standard deviation of the AA scores

$$s_x = \sqrt{\frac{724230 - \frac{(6836)^2}{65}}{65}}$$

$$= 9.0240$$

Standard deviation of the performance data

$$s_y = \sqrt{\frac{533678 - \frac{(5872)^2}{65}}{65}}$$
$$= 7.0280$$

Correct the correlation coefficient for range restriction

$$r_u = \frac{.2354(20/9.024)}{\sqrt{[1 - (.2354)^2][1 - ((20)^2/(9.024)^2)]}}$$
$$= .4730 \text{ (moderate positive correlation)}$$

Test for significance

$$SE = \frac{1}{\sqrt{65-3}}$$
$$= .1270$$

when $r = .4730$ and $N = 65$ the corresponding $z = .51$. The confidence interval at the .05 level of significance for the true z is now $.51 \pm .2489$ or .2611 to .7589. Converting these z 's back into r 's yields a confidence interval from .25 to .64.

The correlation coefficient between the area aptitude scores and the BNCOC scores shows that there is a moderate positive correlation. Since the confidence interval extends from .25 to .64, the true correlation is positive, weak to strong in magnitude.

Straight line equations: $y = a + bx$

$$b = .473[7.028/20]$$
$$= .1662$$

$$a = (5872/65) - [.1662(6836/65)]$$
$$= 72.8593$$

$$y = 72.8593 + .1662x$$
$$= 72.8593 + .1662(100)$$
$$= 89.5 \text{ (predicted cut score)}$$

Frequency Chart

Performance Scores	I-IIIA	IIIB	Overall
65-69	1	0	1
70-74	1	0	1
75-79	1	2	3
80-84	3	5	8
85-89	4	11	15
90-94	8	4	12
95-99	19	6	25
Total	37	28	65

predicted cut score

Conditional Probabilities

	I-IIIA	IIIB	Overall
65	1.00	1.00	1.00
70	.97	1.00	.98
75	.95	1.00	.97
80	.92	.93	.92
85	.84	.75	.80
90	.73	.36	.57
95	.51	.21	.38

According to the table above, if the cut score were set at 90, a I-IIIA would have a 73% chance of being successful, and a IIIB would have a 36% chance of being successful.

III. Research Design

a. Data collection:

(1) The MOS to be studied were placed into groups (09B, 09C, 09D) and (09E, 09F, 09G). Data from the BNCOC end-of-course exam were collected from 1 May-31 Aug on one MOS from each group.

(2) Basic to achieving the study goal, the gathered data were used to determine if a statistical relationship existed between soldier performance and quality as measured by the ASVAB.

b. Analysis:

(1) Alternative hypothesis: Soldiers with higher AFQT scores (I-IIIA) perform better than soldiers with lower AFQT scores (IIIB).

(2) Null hypothesis: Soldiers with higher AFQT scores do not perform better than soldiers with lower AFQT scores.

(3) Descriptive data table

Group	AFQT category	Mean score	Std dev	Sample size (n)
a	I-IIIA	91.84	7.6742	37
b	IIIB	88.36	5.4721	28

(4) Significance level: .05

(5) Rule: Reject the null hypothesis if the obtained t-statistic is greater than the tabled t-statistic.

(6) The formula for determining the t-statistic is:

$$t = \frac{a-b}{sd}$$

substituting previously computed values in the formula yields

$$= \frac{91.84-88.36}{\sqrt{\frac{[(7.6742)^2(37-1)+(5.4721)^2(28-1)][1/37 + 1/28]}{(37+28-2)}}}$$
$$= 2.9098$$

(7) tabled t-statistic: 2.00

(8) Since the obtained t-statistic (2.9098) is greater than the tabled t-statistic, the null hypothesis was rejected. The conclusion was that soldiers with higher AFQT scores perform better than those with lower AFQT scores.

(9) 95% confidence limits

$$\text{Lower limit} = 3.48 - (2.00 \times 1.196)$$
$$= 1.088$$

$$\text{Upper limit} = 3.48 + (2.00 \times 1.196)$$
$$= 5.8720$$

Since both limits are positive, there is a 95% probability that the difference between the two AFQT categories is statistically significant. In other words, it is extremely unlikely that this difference occurred by chance.

c. Descriptive data summary

	Number Tested	Mean BNCOC Score	Std Dev	Mean AA Score	Std Dev
I-IIIA	37	91.84	7.6742	109.05	7.7388
IIIB	28	88.36	5.4721	103.74	7.9753
Overall	65	90.34	7.0280	105.17	9.0240

d. Desired quality accession mix

Step 1. Construct an AA frequency distribution table.

Selector area aptitude composite (GM)

90-94	95-99	100-104	105-109	110-114	115-119	120-124
9	8	13	17	9	7	2

Step 2. Convert AA score to an AFQT distribution.

	90- 94	95- 99	100- 104	105- 109	110- 114	115- 119	120- 124	%
Expected AFQT: distribution :	9	8	13	17	9	7	2	
I-IIIA	3	4	8	12	7	6	2	65
IIIB	3	2	4	4	2	1	0	25
IV	3	2	1	1	0	0	0	10

Per cent by AFQT category*

I-IIIA	IIIB	IV
65	25	10

*desired quality mix

Step 3. Translate desired quality mix at the Target Skill Level into desired accession vector.

I-IIIA	$65/.0645 = 993.88$
IIIB	$25/.0696 = 359.20$
IV	$10/.0837 = 119.47$

$993.88/1472.55 = .67*$
$359.20/1472.55 = .24*$
$119.47/1472.55 = .09*$

*desired accession quality mix

IV. Major Findings

a. The data generally supported the hypothesis that skill level 3 soldiers with higher AFQT scores perform better than soldiers with lower AFQT scores. There is a high probability (95%) that the I-IIIA AFQT category soldiers would come out ahead again if the study were repeated.

b. The sample correlation coefficient provided us a point estimate of the strength of the linear relationship between the performance measure scores and the quality data (ASVAB scores). The coefficient shows a positive correlation that is moderate in magnitude. We can reasonably conclude that as AA scores increase, performance on the test instrument increases. Since AA scores are highly correlated to AFQT scores, soldiers in AFQT Categories I-IIIA will probably perform significantly better than Category IIIB soldiers.

c. Based on the validated performance measure analysis, the level of quality as measured by the ASVAB needed to be accessed is:

I-IIIA: 67%

IIIB: 24%

IV: 9%

Collected data

I-IIIA			IIIB		
AFQT %ile	BNCOC Scores	AA Scores	AFQT %ile	BNCOC Scores	AA Scores
60	91	106	38	89	102
84	94	120	38	81	106
75	88	108	41	94	82
53	95	98	31	94	89
67	66	107	34	83	99
60	74	107	42	87	102
65	81	104	40	86	96
53	98	101	44	79	92
62	96	108	41	82	93
70	99	112	41	92	103
56	99	114	41	88	89
58	97	105	31	85	101
50	96	107	31	84	96
50	93	103	42	97	109
86	95	119	30	96	96
62	98	112	45	96	94
50	87	109	38	95	89
59	89	116	41	97	104
63	91	115	41	96	98
89	84	84	34	79	97
53	77	109	43	85	108
58	84	102	30	86	113
89	91	118	45	87	101
53	85	92	46	84	119
77	93	114	35	85	106
67	90	108	38	88	104
77	90	115	44	91	106
55	95	114	35	88	107
52	99	111			
50	99	114			
58	96	99			
89	98	119			
77	96	113			
84	99	112			
86	98	123			
86	98	108			
80	99	109			

END

8-87

DTIC



AD-A182 387

DISTRIBUTION OF QUALITY PROGRAM HANDBOOK(U) ARMY
SOLDIER SUPPORT CENTER-NATIONAL CAPITAL REGION
ALEXANDRIA VA C ELLIS ET AL. JUN 87

2/2

UNCLASSIFIED

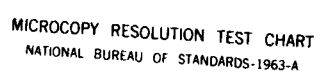
F/G /

NL

END

DATE

18



REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION Unclassified			1b RESTRICTIVE MARKINGS NA		
2a SECURITY CLASSIFICATION AUTHORITY NA			3 DISTRIBUTION/AVAILABILITY OF REPORT Distribution Statement A		
2b DECLASSIFICATION/DOWNGRADING SCHEDULE			5 MONITORING ORGANIZATION REPORT NUMBER(S) NA		
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NA			7a NAME OF MONITORING ORGANIZATION		
6a NAME OF PERFORMING ORGANIZATION US Army Soldier Support Center-- National Capital Region		6b OFFICE SYMBOL (If applicable) ATNC-NMF-A	7b ADDRESS (City, State, and ZIP Code)		
6c ADDRESS (City, State, and ZIP Code) Alexandria, Virginia 22332-0400			9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER		
8a NAME OF FUNDING/SPONSORING ORGANIZATION		8b OFFICE SYMBOL (If applicable)	10. SOURCE OF FUNDING NUMBERS		
8c ADDRESS (City, State, and ZIP Code)		PROGRAM ELEMENT NO.	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
11 TITLE (Include Security Classification) Distribution of Quality Program Handbook (Unclassified)					
12 PERSONAL AUTHOR(S) CPT Clyde Ellis, Mr. Leslie B. Lord, and MAJ Gregory D. Citizen					
13a. TYPE OF REPORT Information		13b. TIME COVERED FROM TO		14. DATE OF REPORT (Year, Month, Day) 8706	
15. PAGE COUNT 56					
16 SUPPLEMENTARY NOTATION NA					
17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
05	09				
19 ABSTRACT (Continue on reverse if necessary and identify by block number) The Distribution of Quality Program Handbook explains what the Distribution of Quality (DQ) Program is, defines the specifications that proponent DQ submissions must meet, and describes one possible method to meet these specifications.					
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a NAME OF RESPONSIBLE INDIVIDUAL Leslie B. Lord			22b TELEPHONE (Include Area Code) 202-325-3703		22c. OFFICE SYMBOL ATNC-NMF-A