AD 45270

UNCLASSIFIED

RESEARCH ON THE DEVELOPMENT OF SHIPBOARD PERFORMANCE MEASURES

TECHNICAL REPORT V

INTERRELATIONSHIPS BETWEEN APTITUDE TEST SCORES, PERFORMANCE IN SUBMARINE SCHOOL, AND SUBSEQUENT PERFORMANCE IN SUBMARINES AS DETERMINED BY RATINGS AND TESTS

Prepared For

Personnel and Training Branch
Office of Naval Research
Department of the Navy

By

Management and Marketing Research Corporation
Los Angeles, California

RESEARCH ON THE DEVELOPMENT OF SHIPBOARD PERFORMANCE MEASURES

TECHNICAL REPORT

V

INTERRELATIONSHIPS BETWEEN APTITUDE TEST SCORES, PERFORMANCE IN SUBMARINE SCHOOL, AND SUBSEQUENT PERFORMANCE IN SUBMARINES AS DETERMINED BY RATINGS AND TESTS

By

Robert R. Mackie, Ph.D.
Clark L. Wilson, Ph.D.
Donald M. Buckner

October 1954

Prepared For

Personnel and Training Branch
Office of Naval Research
Department of the Navy
Washington 25, D. C.

Contractor

Management and Marketing Research Corporation
3004 South Grand Avenue
Los Angeles 7, California

Contracts: N8 onr 70001
Nonr 1241(00)

ABSTRACT

This report describes research that has been conducted to determine the relationships among scores on a variety of aptitude tests, standing in Basic Enlisted Submarine School, New London, and subsequent performance aboard submarines as measured by ratings, written tests, and job sample tests.

The interrelationships of the several shipboard performance measures are described and the results of a factor analysis of the intercorrelations of aptitude test scores and Submarine School criteria are presented.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

We wish to express our sincere thanks for the cooperation of many Naval personnel in making this study possible. Particularly, we wish to thank Captain G. R. Donaho, USN, for permission to begin work afloat in Submarine Flotilla One, U. S. Pacific Fleet. Subsequent Commanders of this Flotilla have been most helpful in the continuing research.

Force, Squadron, Division and Tender staffs in both Pacific and Atlantic Fleets have contributed greatly to the progress of the work and our especial appreciation goes to the Officers and leading Petty Officers aboard the more than thirty submarines where the work was done.

Captain T. L. Willmon, MC, USN, and Commander G. J. Duffner, MC, USN, as Officers-In-Charge, Medical Research Laboratory, U.S. Naval Submarine Base, New London, have both done much to further the work of this project.

We also wish to acknowledge the efforts of Dr. Wayne S. Zimmerman, of San Diego State College, whose work led to a meaningful solution to the factorial analysis described in Chapter V.

Chapter I

SUMMARY AND OPERATIONAL IMPLICATIONS

Research has been conducted to determine the extent to which shipboard criterion measures, developed during earlier phases of this over-all study, were predictable from Submarine School standing and from a variety of aptitude tests administered to enlisted personnel upon entrance into the Basic Submarine School. The shipboard measures included job sample tests, ratings on general traits, ratings on specific technical tasks, and written job knowledge tests.

In addition, the extent to which the Submarine School criterion was predictable from the same aptitude test scores was determined and a tentative identification of some of the basic variables in that criterion was made through a factor analytic study.

From the results of the study, the following conclusions and implications are drawn:

1. Performance in Submarine School is highly predictable from a combination of scores on the Navy GCT, ART, and Mechanical Knowledge tests and a test of the Direction Marking type used in this study.
2. Factor analytic results suggest that there are at least three major factors in the Submarine School criterion that are identifiable from aptitude test results:
    a) Mechanical knowledge or comprehension
    b) Numerical facility
    c) A Grades Factor, which may reflect academic motivation.
3. Measures of shipboard performance are also predictable to a reasonable degree primarily by Mechanical, Numerical, Reasoning, Direction Marking and Visual Attention tests in various combinations.

    Actual tests of performance, such as Job Sample tests, appear

to be more predictable than ratings of abilities to perform specific tasks, ratings on general traits pertaining to job knowledge, or Written Job Knowledge tests.

4. Submarine School standing has a moderate relationship to subsequent shipboard performance, particularly the more technical aspects of that performance. (It was not related to scores on Personal Adjustment traits in the rating scale.)

5. Ratings on general traits using a man-to-man type rating scale have some validity for subsequent shipboard performance as measured by Job Sample tests. This is true of the traits oriented toward the technical competence aspect of performance. It is very likely more true of certain rates aboard ship than of others. In addition, the rating scale yielded a great deal of reliable variance which appeared to be non-technical in nature and which is characterized by a man's attitude toward his job or his adjustment to shipboard life.

6. Ratings on a specific technical check list rating scale have some validity for shipboard performance as measured by Job Sample tests but do not appear to have a clear-cut advantage over more general trait ratings such as Knowledge of the Job and Ability to Trouble-shoot. Raters do not appear to know, with any great certainty, whether or not their men can actually perform certain specific tasks At the lower pay grade levels it is thought that it is often difficult for a rater to isolate the contribution of one man against those of others, particularly when the men typically work in gangs and on complex equipment.

The fact that the specific check list ratings and the more

general trait ratings correlated higher with each other than with any other measure suggests that the check list ratings may not be yielding any more in the way of valid information than would be obtainable from the more general rating scale. An argument against this conclusion, however, is provided by the fact that check list ratings are somewhat more predictable by aptitude test scores than are ratings on general traits.

7. Scores on a Written Job Knowledge test do not appear to be a satisfactory substitute for measures of job skill as revealed in Job Sample test scores. The reason for the relative independence of Written Job Knowledge test scores and Job Sample Performance test scores on the one hand, and of the various sub-tests in a Job Sample Performance test battery on the other, needs further investigation. The relative extents to which general skills and specific knowledges are important for effective performance aboard a submarine was not clear from this study.

8. An appreciable amount of non-technical variance is present in measures of performance aboard submarines when ratings are used. The nature of this variance and its possible predictability would appear to warrant considerable study.

Chapter II

INTRODUCTION TO THE PROBLEM

This report is concerned with the question of the extent to which aptitude tests, administered to trainees at the Basic Enlisted Submarine School, New London, were predictive of subsequent performance aboard submarines as measured by the job sample tests, rating scales and check lists. Additionally, since performance in Submarine School was available as an immediate criterion, the extent to which aptitude test scores predicted it and the extent to which it predicted the shipboard criterion measures also was determined.

The performance of submarine personnel, particularly during time of war, offers a challenge to psychologists and medical specialists who would try to predict behavior which undoubtedly reflects a combination of technical and highly specialized skills on one hand, and personal adjustment to a unique social and physical environment on the other.

The efforts of this project were directed toward the development of suitable measurements of performance aboard submarines which might be used to evaluate the procedures employed, or to be employed, in the screening of enlisted personnel on submarines.

A difficult problem arises in assessing the adequacy of the criterion measures that are developed. The very fact that criteria of performance aboard ship are needed implies that there are no available standards of performance against which the criterion measures that are developed can themselves be evaluated. The ultimate criterion of the performance of submarine personnel is the reduction in striking power of enemy forces. Certainly such a criterion has severe limitations from the point of view of measurement in addition to its general lack of availability. The various intermediate criteria that might be developed require evaluation but elude direct correlation with war effectiveness. In such a situation, and it is typical,

there appear to be at least four questions which should and can be answered in evaluating the relevance of the criterion measures developed:

1. Do operating personnel consider the measures to be relatively complete indices of the really important aspects of performance?
2. Do the measures reflect differences which are known or presumed to be present in personnel of differential experience?
3. Are the measures reliable?
4. Do predictive measures, which are logically related to the criterion measures, actually correlate with them?

Other reports in this series have described the development and measurement characteristics of the several shipboard performance measures developed. These were:

RESEARCH ON THE DEVELOPMENT OF SHIPBOARD PERFORMANCE MEASURES

Part I: The Use Of Practical Performance Tests In The Measurement Of Shipboard Performance Of Enlisted Naval Personnel. (Nov. 1952)

Part II: The Use Of A Performance Rating Scale In The Measurement Of Shipboard Performance Of Enlisted Naval Personnel. (Feb. 1954)

Part III: The Use Of Performance Check Lists In The Measurement Of Shipboard Performance Of Enlisted Naval Personnel. (Feb. 1954)

Part IV: A Comparison Between Rated And Tested Ability To Do Certain Job Tasks. (Feb. 1954)

In these reports information was presented to indicate that the measures

developed did have the characteristics demanded by the first three questions: (1) they were accepted by operating personnel as relevant and important; (2) their scores were highly related to differential experience; (3) they were highly reliable.

The purpose of this report is to answer the fourth, and, from the standpoint of utility, most important question: what was the extent of the correlations between the criterion measures and certain predictors, which, on a logical basis, should be related to them?

Chapter III

SUBJECTS, TESTS AND CRITERIA

Introduction

Beginning in the Fall of 1948 and continuing for over two years, 16 aptitude tests, most of which were well known published tests, were administered to members of each incoming class of the Basic Enlisted Submarine School at New London, Connecticut. Approximately 55% of all incoming personnel took the tests during this period. Testing was accomplished during the days preceding the convening of a Submarine School class. Since some men did not report for duty until the day on which classes actually commenced, not all men in a given class were available for testing. Other than this, there was no selection of subjects, all men arriving at the station prior to commencement of the class being tested.

Testing was accomplished by a research staff member (civilian) and by a Chief Petty Officer attached to the school. The tests were administered to men in groups of up to 90. Testing time for the battery was approximately three hours. A break was given midway during the testing period.

Subjects

The vast majority of subjects were non-rated enlisted personnel who were relatively homogeneous with respect to age and educational level. The age range was from 17 to 30 years with the great majority being 17 to 19. Formal education ranged from 8 to 13 years, with a marked mode at 12 years. Most of the subjects were non-rated personnel, less than 3% of the group being 3rd Class Petty Officers or higher.

The men were informed that the tests were given for research purposes. By and large motivation was considered to have been good with most subjects appearing to enjoy the experience.

The Aptitude Tests

On the pages that follow, the aptitude tests used in the study are described. In addition, mean scores, standard deviations, total number of items, time required for the test and scoring formulae are indicated. It will be noticed that many of the tests were scored in several ways. The following list gives a brief overview of the tests that were employed for the study:

1. The Navy Basic Battery (Scores from these tests were already available for all candidates)

   GCT
   Mechanical
   Mathematical Reasoning
   Clerical

2. Thurstone's Primary Mental Abilities

   Verbal
   Reasoning
   Space
   Number
   Word Fluency

3. Ruch - Wilson Selection System (Form SS)

   Manual Dexterity
   Visual Attention
   Simple Detail
   Eye-Hand Coordination

4. Ruch's Survey of Working Speed and Accuracy

   Coding
   Finger Dexterity
   Accounting

5. Guilford - Zimmerman Aptitude Series

   Part IV, Perceptual Speed
   Part VII, Mechanical Knowledge

6. Unpublished tests from various sources

   Spatial Orientation
   Visual Memory
   Direction Marking
   Social Science

SOCIAL SCIENCES

This is a test of your knowledge of such things as history, government, geography, art, literature, etc.

In this test, there are five possible answers to each item. Only one of these is correct. Decide which is the right answer and put its number in the parenthesis in front of the question.

1. ( ) Which one was a French politician?
   a. Disraeli
   b. Talleyrand
   c. Truman
   d. Henry VIII
   e. Lloyd George

2. ( ) Paraguay is in
   a. Asia
   b. Africa
   c. Northern Europe
   d. The Balkan States
   e. South America

Scoring: 1. Number of items incorrect. Results: Mean: 15 S.D. 6.75
2. Number of items correct. Results: Mean 38 S.D. 11.5
3. Number of items correct minus one-fourth the number of items incorrect. Results: Mean 35 S.D. 12.5

TIME: 30 Min. NO. ITEMS: 72

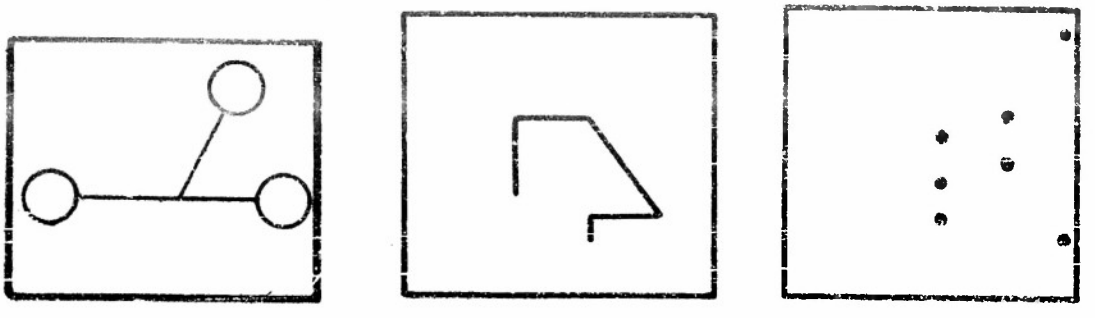
MEMORY FOR DIAGRAMS

This is a test of your ability to remember diagrams.

Your task is to look at a group of ten diagrams for ONE MINUTE and then identify those ten from a group of twenty diagrams on another page.

Study the diagrams carefully.

(Here are shown three diagrams, each of which is from a different group of ten related diagrams.)



Scoring: The number of items right minus the number of items wrong.
Results: Mean 16 S.D. 5

TIME: 1 Min. to memorize, 3 Min. to answer for each part. NO. ITEMS: 30

CODE TRANSLATION

When the examiner says GO, you are to translate a passage which is given in code. The code is made up by assigning numbers to the letters of the alphabet. To translate you look up each number in the key to see what letter it stands for. You then write that letter in the proper space below the number.

THE KEY

1	2	3	4	5	6	7	8	9	10	11	12	13
L	I	A	D	F	Q	T	W	X	N	Z	B	E
14	15	16	17	18	19	20	21	22	23	24	25	26
G	J	M	O	R	U	Y	C	H	K	P	V	S

Work as rapidly as you can without making mistakes. Your score is the number of lines translated without error in five minutes.

The first line in the message is already translated to show how it is done. An * indicates the end of a word or sentence.

(1)	7	22	13	*	24	19	18	24	17	26	13	*	17	5	*	7	13	26
	T	H	E		P	U	R	P	O	S	E		O	F		T	E	S
(2)	7	26	*	25	19	21	22	*3	26	*	7	22	13	*17	10		13	*	7

Scoring: Number of correct letter translations; Results: Mean 102 S.D. 26.5
TIME: 5 Min. NO. ITEMS: 293

PERCEPTUAL SPEED, Form a

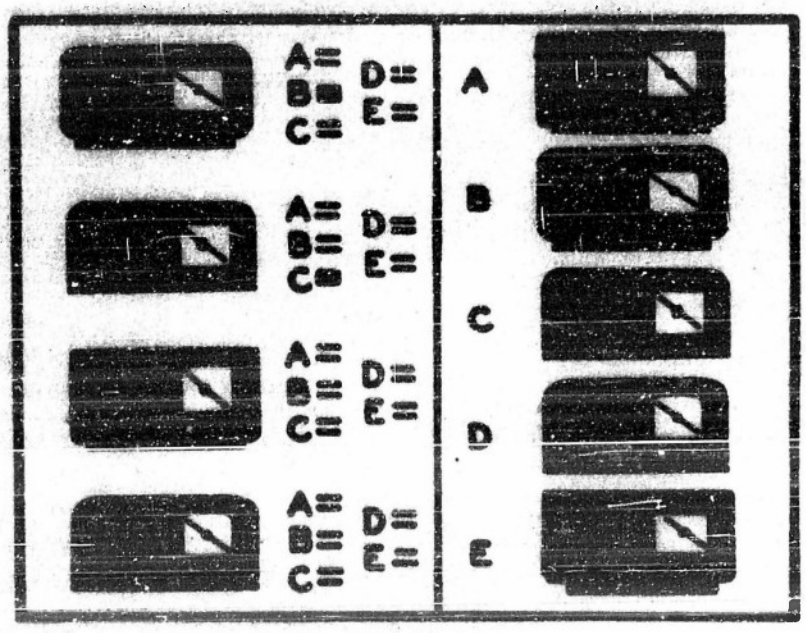
This is a test of how rapidly and accurately you can see objects in order to match them.

Look at the first radio at the left. Which one of the five at the right is most nearly like it? Radio B is the one, so answer space B has been blackened.

Look at the second radio at the left. Which radio at the right is most nearly like it? Radio C is the correct answer, so answer space C is blackened.

Now find the radios most nearly matching the third and fourth ones at the left and blacken the correct answer spaces.

PERCEPTUAL SPEED (CONTINUED)



TIME: 5 Min.
NO. ITEMS: 72

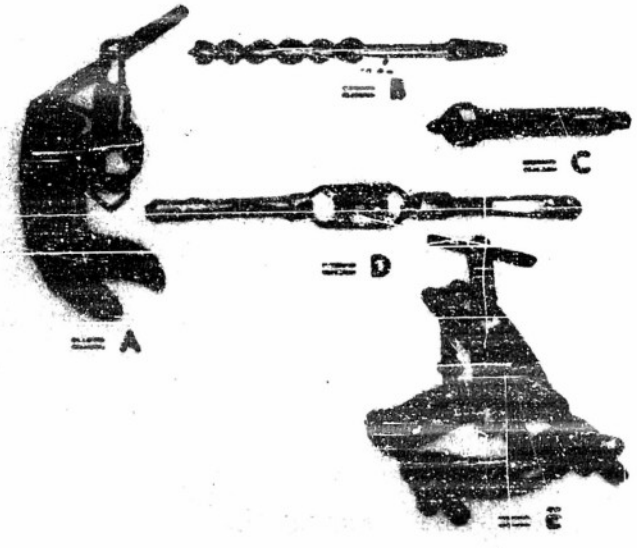
Scoring: 1. Number of items answered correctly minus the number of items answered incorrectly. Results: Mean 42.5 S.D. 8
2. Number of items answered incorrectly. Median 1.25 Q_1 = .5 Q_3 = 2.62
3. Number of items attempted. Results: Mean 45 S.D. 9

MECHANICAL KNOWLEDGE

In this test there are five possible answers to each item, one of which is correct. Decide which is the right answer to each item and blacken clearly the answer space connected with it.

Pick the tool that is best for the described use:

Cut threads



Pick the answer that describes the most likely cause of the trouble:

The distributor points, lights, coil, etc., burn out every few weeks.

A. Overcharged battery
B. Undercharged battery
C. Direct short in the system
D. Loose battery cables
E. Oversized ignition wires

Scoring: Number of correct items.
Results: Mean 22 S.D. 8.5
TIME: 15 Min. NO. ITEMS: First 47

VERBAL-MEANING

The first word in the following line is BIG.

BIG A. Ill B. Large C. Down d. Sour

One of the other words means the same as BIG. This word is Large. Large is answer B. An X has been marked in B on the Answer Pad.

Scoring: Number of items correct. Mean 29 S.D. 4.5
TIME: 6 Min. NO. ITEMS: 50

SPATIAL ORIENTATION, Form b

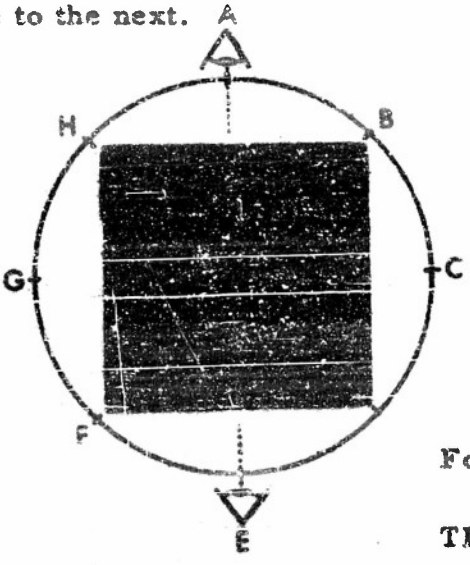
This is a test of your ability to find your position on a circle by looking at the relative position of objects inside the circle.

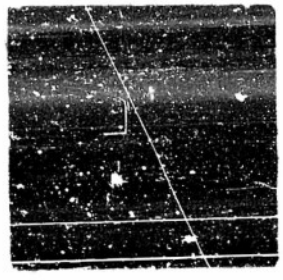
Look at the pictures below. At the left side of the page inside the circle, is a TOP VIEW picture of our balls in a line. This picture was taken from a point directly above the balls. That is, you are looking straight down on the formation.

To the right of this picture are three other pictures of the same formation of four balls. THE BALLS HAVE NOT BEEN MOVED. The pictures are taken from a position off to the side of the formation of balls. In all of them you are looking down at an angle of about forty-five degrees.

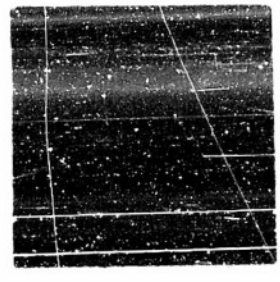
Your task is to decide where you would be standing ON THE CIRCLE to get these pictures.

For example, in the first picture the broad stripe is running straight away from you and the balls are in a straight line. You must be standing at either "A" or "E". The single small white ball is toward you and the two small white balls are on the other side of the big ball from you. Hence you must be standing at "E". You will see an eye drawn on the paper at "E". Put the letter "E" in the box under the first picture. NOTE: The shadows in the pictures are not to be used as cues because they move from one picture to the next.





1. ☐



2. ☐

Form b: Scoring: Number of correct items.

Results: Mean 35 S.D. 9
TIME: 6 Min. NO. ITEMS: 45

SPACE

Look at the row of figures below. The first figure is like the letter F. All the other figures are like the first one, but they have been turned in different directions.

Now look at the next row of figures. The first figure is like the letter F. But none of the other figures look like an F, even if they were turned right side up. They are all made backward.

Some of the figures in the next row are like the first figure. Some are made backward.

In the row of figures below, mark an X in the box of EVERY figure which is LIKE the first figure. Do NOT mark the figures which are made backward.

Scoring:

1. Number of items correct.. Mean 29 S.D. 8.5
2. Number of items incorrect. Median 1.2 Q_1 = .50 Q_3 = 3.0
3. Number of items correct minus the number of items incorrect Mean 27 S.D. 9.5

TIME: 5 Min. NO. ITEMS: 54

REASONING

Study the series of letters below. What letter should come NEXT ?

a b a b a b a b

The series goes like this: ab ab ab ab. The NEXT letter in the series should be a. An X has been marked in [a] on the Answer Pad.

Now study the series of letters below. Decide what the NEXT letter should be. Mark an X in the box of the NEXT letter in this series.

c a d a e a f a

This series goes like this: ca da ea fa. You should have marked g.

Scoring: 1. Number of items correct. Mean 16.5 S.D. 5.5

TIME: 5 Min. NO. ITEMS: 59

NUMBER

At the right are two columns of numbers which have been added. Add the numbers for yourself to see if the answers are correct.

	A	B
•	16	42
	38	61
	45	63
	99	176

The A answer is Right, so an X has been marked in [R] on the Answer Pad.

The B answer is Wrong, so an X has been marked in [W] on the Answer Pad.

Now check the sums of the problems below. If the answer is Right, put an X in [R]. If the answer is Wrong, put an X in [W]. If you wish to change an answer, draw a circle around this box like (X). Then mark the new answer in the usual way.

A	B	C
17	35	63
84	28	17
29	61	89
140	124	169

Scoring: 1. Number of items correct. Mean 26 S.D. 8
2. Number of items incorrect. Median 1.5 Q_1 = .50 Q_3 = 2.80
3. Number of items correct minus the number of items incorrect. Mean 24 S.D. 9

TIME: 5 Min. NO. ITEMS: 70

WORD-FLUENCY

Look at the words in the list below. Each word begins with *d*.

doll
dimes
daisy
doughnut

You are to write several words which BEGIN with *p*. One word you might write is *potty*. Turn to the next page, and in the spaces at the TOP of the page on the Answer Pad, write three more words which BEGIN with *p*.

Scoring: 1. Number of correct words. Mean 40 S.D. 10

TIME: 5 Min. NO. ITEMS: SPEEDED

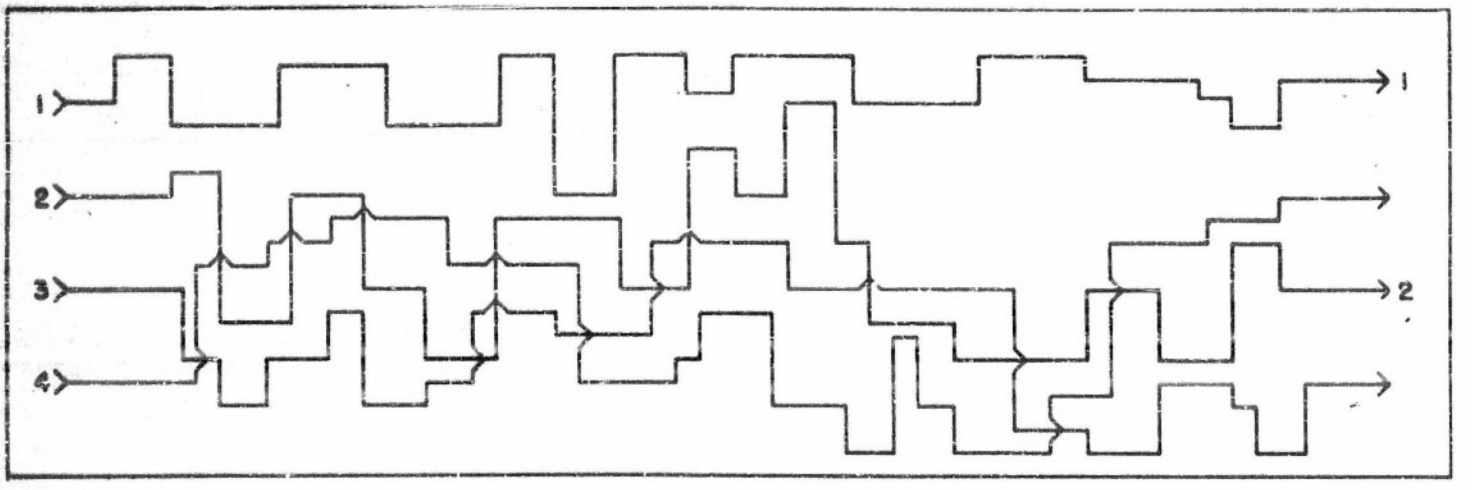
VISUAL ATTENTION

This is a test of your ability to keep your eyes on one object as you sweep through a field of many similar objects.

Your task is to follow each line from the left side of the page where it begins across to the right side of the page where it ends. You will then write the number of the line by the arrow head to show your answer.

For example, in the sample below, lines No. 1 and 2 have already been done to show you how. Go ahead and do 3 and 4 for practice. Remember to write the number of the line where that line comes out on the right.

Do not use your pencil to follow the line if you can help it. You will work faster if you depend on your eyes alone.



Scoring: 1. Number of items RIGHT; Results: Mean 16 S.D. 4.25

2. Number of items WRONG; Results: Median .5 Q_3 = 1.10

TIME: 5 Min. NO. ITEMS: 40

MANUAL DEXTERITY

This is a test of your ability to make rapid movements with your hands and fingers.

Your task is to make as many "gates" as you can in the time allowed. Accuracy is not important. However, you must have four vertical strokes and one diagonal stroke--that is--five strokes in all.

Below is a space for you to practice to be sure you know what you are to do. It makes no difference which way the diagonal runs: left to right or right to left.

(A) 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸

(B) _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

Scoring: Number of "gates" completed

Results: Mean: 44 S.D. 7.5

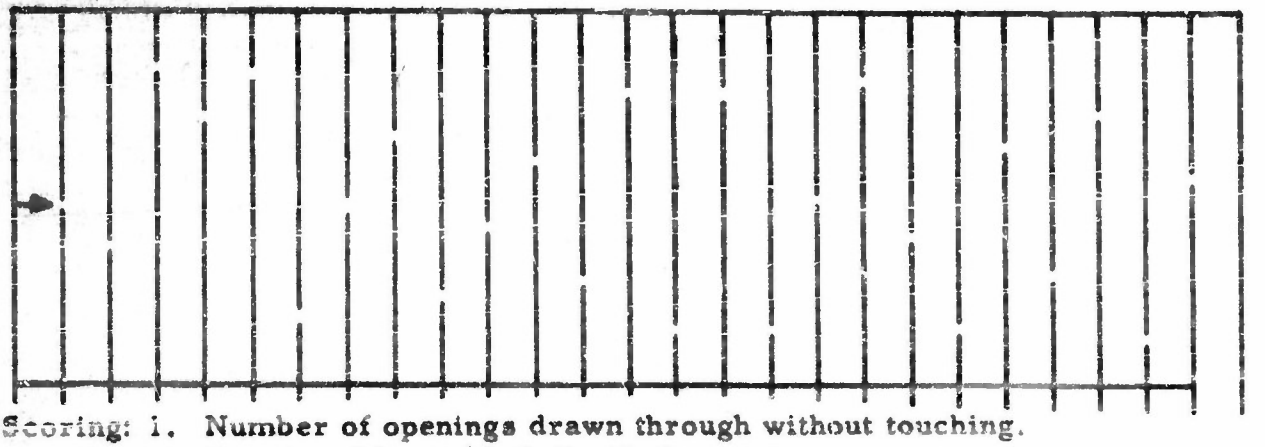
TIME: 1 Min. NO. ITEMS: SPEEDED

EYE-HAND COORDINATION

This is a test of your ability to coordinate your eye and hand in a continuing task.

Your task is to trace a line through the openings in the vertical lines starting at the arrow in the upper left hand corner. Draw a continuous line across the page to the right going through the openings in the vertical lines. When you come to the last alley on the right then go down that alley to the next opening and start back to the left. BE CAREFUL NOT TO TOUCH THE VERTICAL LINES AS YOU GO THROUGH THE OPENINGS.

Below is a sample problem for you to work on.



Scoring: 1. Number of openings drawn through without touching.
Results: Mean: 68 S.D. 15.5
2. Number of errors. Results: Mean 22.5 S.D. 19
TIME: 2 Min. NO. ITEMS: 145

RECOGNITION OF SIMPLE DETAIL

This is a test of your ability to detect a simple expected detail in your visual field.

Your task is to put a check mark at the end of each line in which all letters are "Q's". If there is an "O" in the line any place, put an "X" in the blank at the end of the line.

The samples below have been marked to show you how. You work the unmarked samples to give yourself some practice.

(A) QQQQQQQQQQQQQQQQQQQQ ✓
(B) QQQQQQQQQQQQQQQQQQQQ X

Scoring: 1. Number of total items attempted; Results: Mean 111 S.D. 22
2. Number of items RIGHT: Results: Mean 105 S.D. 21
3. Number of items WRONG: Results: Mean 4 S.D. 3.75
TIME: 4 Min. NO. ITEMS: 150

FINGER DEXTERITY

When the examiner says GO, put a pencil dot in each of the O's working from left to right or from right to left. Work as rapidly and as accurately as you can. The time limit is five minutes. Dots must not touch O's.

OOOOOOOOOOOOOOOOOOOO (1)
OOOOOOOOOOOOOOOOOOOO (2)

Scoring:
1. Number of O dotting attempts rounded to the nearest interval of five. Results: Mean 450 S.D. 93.75
2. Number of O dottings in which the pencil mark is located outside the O. In this case a single miss is sufficient to invalidate the entire line of twenty attempts. Where less than twenty attempts are made in a line and a miss is made, then the number of misses is rounded to the nearest multiple of five. Results: Median 1.19 Q_1 = .33 Q_3 = 3.00
3. Number of O dottings in which the pencil mark is located inside the O. All attempts in each line must be correct for credit. Results: Mean 310 S.D. 140
4. Test is divided into three parts. The number of O dottings that meet the criteria of 3 above in part three is subtracted from part one. Results: Mean 20 S.D. 40

TIME: 1 1/2 Min. for each part NO. ITEMS: SPEEDED

COUNTING

This task consists of counting each of the vowels in each line. You will notice that there are forty lines of text below. At the end of each line there are five blanks headed A, E, I, O, and U. For each line you are to count and record the number of each of those letters in the line. Count all five vowels and record them for each line before going on to the next line. If a particular vowel does not occur in a given line, record an O in the appropriate blank at the end of the line. Take the lines in order. Count one vowel at a time.

Line (1) is done to show you what is required. Work as rapidly and as accurately as you can. You will have five minutes to do as much of this task as you can.

(1) The statement that all men are created equal
A's (6) E's (8) I's (0) O's (0) U's (1)

(2) in all things is not true. This fact about human
A's ( ) E's ( ) I's ( ) O's ( ) U's ( )

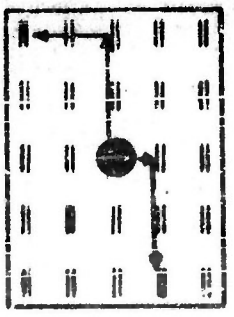
Scoring:
1. Number of attempts made. Results: Mean 47 S.D. 8.5
2. Number of correct responses. Results: Mean 37.5 S.D. 8.25
3. Number of incorrect responses. Results: Mean 9 S.D. 4.5

TIME: 5 Min. NO. ITEMS: 200

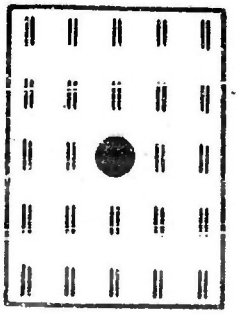
DIRECTION MARKING

This is a test of your ability to think rapidly and accurately about directions. Look at the sample items that follow:

1. Near right, far down
2. Far up, far left
3. Near down, near left
4. Far right, near up



5. Far down, near right
6. Near up, far right
7. Far left, far up
8. Near left, near down



Your task is to place a mark at the distance and direction from the dot in the center of the box. The meanings of the instructions are as follows:

1. Near right means one step to the right from the center dot; near left; one step to the left.
2. Far right means two steps from the dot to the right; far left, two steps to the left.
3. Near up means one step up from the dot; near down, one step down.
4. Far up means two steps up from the dot; far down, two steps down.

Now, you work problems 5 to 8 in the sample item to the right above in the same manner. Place your marks in the appropriate spaces.

Scoring: 1. Number of items incorrectly marked. Results: Median 1.10 $Q_1 = .47$ $Q_3 = 2.25$

2. Number of items correctly marked minus one-fourth the number of items incorrectly marked for the first 37 items. Results: Mean 34.5 S.D. 4.75
3. Number of items correctly marked minus one-fourth the number of items incorrectly marked for 48 items. Results: Mean 38 S.D. 9

TIME: 5 Min. NO. ITEMS: 48

The Criterion of Submarine School Performance

The criterion of Submarine School performance was final standing in class. (Submarine School classes typically consisted of between 100 and 200 men at the time of this study.) Final standing was determined from a man's final average which in turn was a composite score developed from three sources:

1. Aptitude for submarines: This was a rating made by each trainee's section instructor pertaining to the likelihood of his becoming a successful submariner. Only 3 numerical score values were used: 2.75, 3.25 and 3.75. (It will be remembered that a mark of 4.0 is regarded as reflecting the best possible performance according to Navy standards.)
2. Operating average: This is a composite of ratings made by several instructors who teach the student. It is an assessment of the trainee's performance in various training devices such as that calling for the operation of bow and stern planes. The numerical scores assigned to each trainee were limited to the same values used in rating "Aptitude for Submarines".
3. Examination average: This is an unweighted arithmetic average of 5 written examinations given to each trainee during the 8-week course. A typical examination of this kind contained approximately 120 questions. About 50% of the questions were of the multiple choice variety, four alternative answers being offered. The remaining questions required the subject to identify from sketches various components of submarine operating gear. Several alternatives were given and the trainee was asked to match the correct answer with the numbered component.

The Special Case of Subjects Eliminated from School During Basic Training

In order to ascertain the full effectiveness of the aptitude tests for predicting performance at the Submarine School, it was thought that the performance of certain trainees who were eliminated during the course of basic training should be taken into account. Nine different criteria have been stipulated for the elimination of a candidate from submarine training. These are:

1. Lack of ability to do academic work of specialty.
2. Lack of ability to do practical work of specialty.
3. Lack of application.
4. Unwillingness to do work assigned.
5. Temperamentally unfit for submarine duty.
6. Temperamentally unfit due to lack of desire for submarines.
7. Physically unfit for submarine duty.
8. Disciplinary problem.
9. Psychologically not adapted for submarine training.

It was decided that with two exceptions, those physically disqualified and those disqualified for disciplinary reasons, elimineee should be included in the prediction study. (It appeared unreasonable to predict physical fitness by means of aptitude tests and the number disqualified for disciplinary reasons was too small to be predicted reliably.)

Since final standing in Submarine School would not be available for the disqualified students, it was assumed for these subjects that their elimination from training should be regarded as placing them in the lower half of their class. Because tetrachoric correlations were used in the analyses described later, it was not necessary to make a more refined guess as to where these disqualified persons logically should have been placed. The number of eliminees included in the sample of trainees studied accounted for 8.4% of the total number of cases.

Reliability of the Submarine School Criterion

In order to determine whether the reliability of the criterion would be a seriously limiting factor on correlations with aptitude test scores, grades obtained by trainees during the fourth week of school were correlated with those obtained during the eighth week.

For this purpose a sample of 100 men was obtained by taking every fourth man from six different classes. Pearson product-moment correlations were then computed between the fourth and eighth weeks' scores and the results corrected by the Spearman-Brown formula. The results obtained appear in Table I. A second sample drawn from five later classes yielded highly similar results.

TABLE I

THE RELIABILITY OF SUBMARINE SCHOOL CRITERIA
BASED ON A SAMPLE OF 100 MEN SELECTED FROM
CLASSES 40 TO 45
(Fourth week vs. Eighth week scores)

Measure	N	r (Corrected)
Aptitude for submarines	100	.73
Operating average	100	.48
Examination average	100	.53

An estimate of the reliability of the composite criterion, on which final standing was based, was .90. It was concluded that the Submarine School criterion was sufficiently reliable so as to not seriously limit any relationships with predictors.

Shipboard Criteria

Three measures of performance aboard ship were developed for this study. Their development was oriented toward measurement of two broad aspects of performance: (1) that which could be regarded as primarily technical in nature; and (2) that which could be regarded as mainly adjustive or attitudinal toward shipboard and Navy life. The three devices for assessing shipboard performance and

their measurement characteristics are described in detail in Parts I, II and III of this series of reports. They will be mentioned only briefly here.

1. Performance rating scale. A graphic man-to-man performance rating scale consisting of 10 traits was completed by three of each man's superior officers as a part of the shipboard criterion. The raters usually were the two leading Chief Petty Officers of the gang and the division officer of the man concerned. Some of the traits referred to a man's technical performance aboard ship and some were more concerned with attitudinal and adjustive behavior. The performance referred to in the rating scale was described in behavioral terms which were as concrete as possible but necessarily quite general. The rate-rerate reliability and inter-rater agreement on this scale proved to be highly satisfactory, the former being above .80 for periods of from five to nine months and the latter ranging from .45 to .65 for single traits and better than .70 for total score. (See Part II, this series of reports.[1])

2. Performance check lists. Two performance check lists were developed for this study, one for Electrician's Mates and one for Enginemen. These also were rating devices to be filled out by each man's leading petty officer and division officer but, unlike the rating scale, they contained very specific statements about technical job performance. They were designed to reflect as many specific bits of regular important technical shipboard behavior, characteristic of the two rates, as possible. The inter-rater agreement on total scores on

[1] Wilson, C.L., Mackie, R.R. and Buckner, D.N., RESEARCH ON THE DEVELOPMENT OF SHIPBOARD PERFORMANCE MEASURES. PART II "The Use of A Performance Rating Scale In The Measurement of Performance of Enlisted Naval Personnel" (NR-153-625) Management and Marketing Research Corporation, February 1954.

these check lists was very high, above .80. (See Part III, this series of reports.[2])

3. Job sample performance tests. The third shipboard criterion measure was comprised of batteries of job sample tests designed around the Electrician's Mates and Enginemen jobs. A sample of shipboard tasks which was as representative of the job as possible, under the limits of testing time and operational conditions, was developed into a battery which could be administered under standard conditions and objectively scored. Also included in the battery of performance tests was a written job knowledge test of the conventional multiple response, objective type, and a test of knowledge of safety precautions. Results of the administration of these tests showed that individual tests had moderate to high reliability (.50 to .80) and that the composite battery for both Electrician's Mates and Enginemen had high reliabilities (.90). (See Part I, this series of reports.[3])

---

[2] Wilson, C.L., Mackie, R.R., and Buckner, D.N., RESEARCH ON THE DEVELOPMENT OF SHIPBOARD PERFORMANCE MEASURES. PART III "The Use of Performance Check Lists In The Measurement of Shipboard Performance of Enlisted Naval Personnel" (NR-153-625) Management and Marketing Research Corporation, February 1954.

[3] Wilson, C.L. and Mackie, R.R., RESEARCH ON THE DEVELOPMENT OF SHIPBOARD PERFORMANCE MEASURES. PART I "The Use of Practical Performance Tests In The Measurement of Shipboard Performance of Enlisted Naval Personnel". (NR-153-625) Management and Marketing Research Corporation, February 1954.

Chapter IV

VALIDITY OF APTITUDE TESTS FOR THE SUBMARINE SCHOOL CRITERION

Preliminary to the analysis of the relationships discussed in this and subsequent chapters, raw scores on all tests and criterion measures were converted into standard form. In order to facilitate correlational procedures, all standard scores in turn were converted into positive whole numbers, ranging from 0 to 9 (the STEN scale). STEN scores of 0 through 4 included the lower half of the scores for any particular variable, while STENS of 5 and higher included the upper half.

For the purposes of this study, it was considered desirable to be able to combine students from several Submarine School classes on whom predictor scores were available. The over-all plan was to determine the validities of the several aptitude tests on one-half of the total available sample and cross-validate these results on the other half. In order that these samples be of appreciable size, it was necessary to combine men from different submarine classes into larger groups. This procedure also was thought to remove any bias in the results that might be associated with particular classes or with recruits at particular times of the year. Before such a combining procedure could be initiated, however, it was necessary to determine whether or not there were systematic differences between Submarine School classes with respect to the general levels of aptitude involved.

An initial hurdle in the screening of Submarine School candidates was that they must have had a combined score of 105 on the two Navy basic battery tests, GCT and ARI. In order to determine whether samples from the several classes were similar with respect to scores on these two tests, means of the unweighted sum of GCT and ARI scores were computed for each of the nine classes used. The significance of the differences between all possible combinations of these mean scores was tested by means of the t-test. Using the 5% level of confidence as the criterion for rejecting the null hypothesis, the t-tests showed in all cases that there

were no significant differences between the several classes in the abilities measured by the tests. In Table II, the means, standard deviations and standard errors may be inspected and in Teble III, a matrix of the t-values between classes can be seen.

TABLE II

MEANS, STANDARD DEVIATIONS AND STANDARD ERRORS OF THE MEAN OF COMBINED G.C.T. AND A.R.I. SCORES FOR SEVERAL SUBMARINE SCHOOL CLASSES

Class	N	Mean	Standard Deviation	Standard Error of the Mean
40	93	115.01	12.72	1.33
41	121	113.77	11.97	1.09
42	95	114.74	10.35	1.07
43	80	116.06	11.30	1.27
44	73	115.97	11.05	1.30
45	113	114.83	11.00	1.04
46	103	114.33	10.84	1.07
47	144	113.71	10.05	0.88
48	128	114.85	10.28	0.91
Total	950	114.70	11.05	

TABLE III

T-RATIOS FOR ALL MEAN DIFFERENCES OF COMBINED GENERAL CLASSIFICATION AND ARITHMETICAL REASONING TEST SCORES FOR SEVERAL SUBMARINE SCHOOL CLASSES

Class	40	41	42	43	44	45	46	47	48
40	--								
41	.721								
42	.158	- .635							
43	-.571	-1.368	-.795						
44	.516	-1.296	-.730	.050					
45	.107	- .703	-.060	.750	.685				
46	.398	- .367	.271	1.042	.974	.336			
47	.690	- .100	.599	1.392	1.312	.675	.303		
48	.099	- .761	-.078	.775	.706	-.014	-.370	-.742	

With these results, men from all classes were then combined to form a total sample of 950 cases on whom test scores were available as predictors. The men were then assigned alternately to the validation and cross-validation groups, making a total of about 475 men in each sub-sample.

An Earlier Study of the Validity of Tests for Submarine School Candidates

A report by Bartlett, published in 1945, describes an earlier validation of procedures for the selection of Submarine School candidates.[1] Bartlett presented findings of validities for the several Navy basic battery tests using Navy grades based on operations and written examinations, as the criterion. The description of the criterion in his study suggests that it was similar to, but not identical

[1] Bartlett, N.R., "Correlations of Tests With Grades In Submarine School", Bureau of Medicine and Surgery Research Project No. X-243, Medical Research Department, U.S. Submarine Base, New London, Connecticut, February 1945.

with, the one used in this study. With a sample of one class totaling 210 subjects, the product-moment correlation coefficients for various aptitude tests were as follows:

General Classification Test	.25
Arithmetical Reasoning Test	.30
Mechanical Knowledge (Electrical)	.26
Mechanical Knowledge (Mechanical)	.35
Reading	.20
Mechanical Aptitude Test	.23
GCT + ARI	.31
GCT + Mechanical Knowledge (Electrical)	.35

From these findings, Bartlett concluded that:

> "A combination of scores for the General Classification Test, Arithmetic Reasoning Test and Mechanical Knowledge Tests is a better index of grades in Submarine School than is the General Classification Test score alone."

The coefficients reported by Bartlett are all underestimates of the predictive validity of the tests since they are uncorrected for restriction in range. It might be added that they would also be underestimates if only the passing candidates were considered in the study. It is not known whether this was the case in Bartlett's study.

Validity of Single Tests in the Present Study

In Table IV, the tests in this study which proved to have validities significantly greater than zero for predicting Submarine School standing are listed. It will be remembered that candidates who were eliminated from training for one reason or another were assigned to the lower half of their Submarine School class in

computing these coefficients.

TABLE IV

VALIDITY OF SEVERAL TESTS FOR THE SUBMARINE SCHOOL CRITERION
(Tetrachoric Coefficients)
N = 475*

Direction Marking	.46
Navy Mechanical	.43
Navy GCT	.39
Navy ARI	.34
Guilford-Zimmerman Mechanical Knowledge	.32
Thurstone Space	.26
Social Sciences	.24
Thurstone Reasoning	.23
Spatial Orientation	.18
Finger Dexterity	.17
Thurstone Number	.15

* Except as noted in the text.

In the Appendix, the entire matrix of correlations between Submarine School criteria and aptitude test scores, including those for more than one kind of score for many of the tests, can be inspected. For the most part, these coefficients are based on 475 cases, one-half of the number of men in the entire study. For two of the variables, the Direction Marking test and the Social Science test, N was 319 since these tests were not developed when testing first began in New London.

Validities of the Aptitude Test Scores in Combination

Using the Gengerelli approximation, multiple correlation coefficients next were computed between various combinations of test scores and the school criterion. In Table V, the values obtained by this procedure may be inspected.

TABLE V

VALIDITIES, BETA WEIGHTS, AND THE MULTIPLE CORRELATION COEFFICIENTS FOR SEVERAL TESTS AND THEIR COMBINATIONS

Test Number	Test	Validity Coeffi-cient	Beta Weight	Test Battery	Multiple R
50	Direction Marking	.46	.46	50	.46
12	Mechanical Aptitude	.43	.34	50+12	.58
62	General Class. + Arith. Reas.	.40	.10	50+12+62	.59
11	General Classifi-cation	.39	.11	50+12+11	.59
13	Arithmetical Reasoning	.34	.19	50+12+13	.61

The Direction Marking test, which showed the highest validity for the Submarine School criterion, was the first choice by the Gengerelli technique. When the Guilford Zimmerman Mechanical Aptitude test was added to the battery, the multiple R became .58. A combination of GCT + ARI had the next highest validity and, when added, increased the multiple to only .59. Substituting GCT for GCT + ARI also yielded a multiple R of .59. Finally, it was found that the addition of ARI to the first two tests increased the multiple to .61. This appeared to be the most predictive, yet economical, combination of aptitude tests for the Submarine School criterion in this study.

Cross-Validation

Having obtained a multiple R as large as this, especially when compared to that which would have obtained through the regular selection battery (GCT+ARI) for the same subjects (.40), the question of cross-validation then immediately arose. Since only 516 subjects had taken all three of the selected tests, the cross-validation study was limited to that number.

Using the appropriate beta weights, a criterion score was predicted for each man in the cross-validation sample from his scores on the Direction Marking, Guilford-Zimmerman Mechanical Knowledge, and Navy ARI tests. These predicted scores then were correlated with actual class standing. The coefficient of correlation obtained was R = .55, indicating that shrinkage had not been severe and that this battery of tests represented a substantial improvement over the screening battery then in use.

Correction for Restriction in Range

In order to appreciate the full selective power of aptitude tests as predictors of Submarine School performance, it was necessary to determine the validity of the test battery without the restriction in range introduced by selection of candidates on the basis of a combined ARI and GCT score of 105. Correcting for this restriction should reveal the results that would have been obtained had the submarine trainees been representative of the entire recruit population with respect to the abilities measured by these two tests.

Correction for restriction in range is a function of the intercorrelations of the tests and their variances. It was first necessary to obtain a satisfactory estimate of these statistics for the two tests concerned. Stuit[2] has reported inter-correlations between GCT and ARI of .69, .79 and .72 on random samples of 500, 933 and 803 cases. Using these values, an average of .74 was obtained after using Fisher's Z-transformation.

Since the scores for both tests were in standard form, with a population mean of 50 and a standard deviation of 10 by design, it was possible to use 10 as an estimate of the variability in an unrestricted population. There is considerable evidence from Stuit and other sources that this value holds up reasonably well

---

[2] Stuit, D.B., (Ed.) Personnel Research and Test Development, Princeton University Press, Princeton, New Jersey, 1947.

empirically. The variance of the unweighted composite scores in the two tests was then obtained by using the formula:[3]

$$\sigma_s^2 = \sigma_1^2 + \sigma_2^2 + 2r_{12}\sigma_1\sigma_2 .$$

Using this equation and the estimates of population parameters described above, the estimate of the variance in the population without restriction in range was 347.82.

The next task was to estimate the variance for the restricted sample, that is, the sample of Submarine School candidates. To do this, the unweighted scores from the two tests, GCT and ARI, were added since this is the procedure that was actually followed in establishing the cut-off score. The variance of combined scores was then computed on the restricted group of 950 trainees and found to be 122.10. This indicated that rather severe restriction in range had occurred as a result of selection procedures. The ratio of the variance in the unrestricted recruit population to that in the Submarine School candidate population was 2.85.

In order to determine the multiple correlation coefficient for the experimental test battery for an unrestricted group, it was necessary to correct not only the correlations of each test with the criterion for restriction in GCT + ARI scores, but the inter-correlations of the tests as well. Two formulae were necessary to accomplish this task.[4] Formula (1) below was employed to determine the corrected validity coefficients between each test variable and the variable of ARI + GCT on which the restrictions took place.

Formula (2) was employed to determine the corrected coefficients between two

---

[3] Guilford, J.P., Fundamental Statistics in Psychology and Education. McGraw-Hill Book Company, Inc., New York, Toronto, London, 1950. p. 454.

[4] From Thorndike, R.L., Personnel Selection, New York: John Wiley and Sons, 1949, pp. 173-174.

test variables which were both subject to restriction by a third variable (ARI + GCT). These were the corrected intercorrelations of the predictor variables.

$$(1) \quad R_{12} = \frac{r_{12}\,\frac{S_1}{s_1}}{\sqrt{1 - r_{12}^2 + r_{12}^2\,\frac{S_1^2}{s_1^2}}}$$

(Where S and s are the standard deviations in the unrestricted and restricted populations respectively.)

$$(2) \quad R_{12} = \frac{r_{12} + r_{13}\,r_{23}\left(\frac{S_3^2}{s_3^2} - 1\right)}{\sqrt{\left[1 + r_{13}^2\left(\frac{S_3^2}{s_3^2} - 1\right)\right]\left[1 + r_{23}^2\left(\frac{S_3^2}{s_3^2} - 1\right)\right]}}$$

Using the corrected validities and test intercorrelations obtained in this manner, the Gengerelli approximation again was used to estimate the multiple R for the unrestricted population. When this was done, R was found to be .83. A similar procedure using the regular screening battery (GCT + ARI) yielded a corrected R of .59. T-tests of the difference between coefficients resulting from use of the Navy and experimental test batteries were significant beyond the 1% level of confidence in favor of the experimental battery under both restricted and unrestricted conditions. These results are summarized in Table VI.

TABLE VI

COMPARISON OF NAVY AND EXPERIMENTAL TEST BATTERIES FOR PREDICTING THE SUBMARINE SCHOOL CRITERION

Battery	Test Number	Multiple R*	t	Significance	Multiple R**	t	Significance	N
Navy	62	.40			.59			475
			2.74	<.01		6.99	<.01	
Experimental	50, 12, 13	.55			.83			316

* Uncorrected for restriction in range.
** Corrected for restriction in range.

Chapter V

FACTORIAL ANALYSIS OF THE EXPERIMENTAL APTITUDE TESTS AND SUBMARINE SCHOOL CRITERIA

In an effort to identify the basic composition of the experimental test battery and, if possible, learn something about the underlying nature of the Submarine School criterion, a centroid factor analysis of 27 variables including Final Average and its three components was carried out.

For the most part this analysis was quite successful, several well-identified factors clearly emerging. In addition, some other meaningful factors were tentatively identified but these should be verified in other studies before their nature can be regarded as firmly established. It is felt that some important characteristics of the Submarine School criterion were uncovered.

Abbreviated tables of loadings on the several factors, together with tentative identifications of each factor, appear on the following pages.

The complete tables of loadings, before and after orthogonal rotation, appear in the Appendix.

FACTOR I: MECHANICAL

Navy Mechanical	.83
Guilford-Zimmerman Mechanical Knowledge	.70
APT for S/M	.65
Final Average	.64
Operating Average	.58
Examination Average	.53
GCT	.43
Space - O	.36
Thurstone Space	.34

Factor I is clearly mechanical in nature, resembling closely the Mechanical Experience factor found in many analyses. It is believed to involve knowledge of mechanical apparatus and principles rather than innate ability to be a mechanic.

The Submarine School criteria are heavily loaded on this factor. This is regarded as an important, though not unexpected, finding of the study. The

analyses are consistent in reflecting the importance of screening Submarine School candidates on mechanical and spatial tests.

FACTOR II: PERCEPTUAL SPEED

Guilford-Zimmerman Perceptual Speed	.56
Counting	.51
Gate Making	.50
Coding	.49
Simple Detail	.32
Line Following	.27

Factor II clearly appears to be the well-known Perceptual Speed factor. It is described by French[1] as characterized by the task of finding in a mass of distracting material a given configuration which is borne in mind during the search. The fact that the Gate Making test appears heavily loaded on this factor may reflect a lack of other tests in the battery calling for Finger Dexterity.

FACTOR III: VERBAL-ACADEMIC

Social Science	.79
GCT	.70
Social Science Bluffs (lack of bluffing)	.69
Thurstone Verbal	.56
Thurstone Word Fluency	.36

Factor III is similar to that usually defined by social studies or social science tests. It no doubt is confounded with the verbal factor in this analysis, necessitating a complex designation. The word fluency test also comes out on this factor no doubt because there were no other variables in the analysis to help define the fluency factor.

The failure of the Submarine School Final Examination and Final Average to come out on this factor is readily explained by the prior screening of the

[1] This and other definitions have been borrowed from French, J.W., "The Description of Aptitude and Achievement Tests in Terms of Rotated Factors" Psychometric Monographs, Number 5, 1951. p. 278.

candidates' verbal ability on the basis of GCT and ARI scores.

FACTOR IV: GRADES-MOTIVATION

Final Average	.53
Examination Average	.51
O-Dotting	.51
Direction Marking	.40
Operating Average	.39
Tracing	.39
Gate Making	.35
Aptitude for S/M	.32

A number of studies have produced a Grades Factor on which there were insufficient loadings from well identified tests to determine anything about its psychological nature. Among its possible interpretations, motivation, academic interest and persistence have been advanced. Factor IV suggests itself as such a factor because of the high loadings of Final Average and the presence of the other test variables which may possibly reflect motivation or persistence. Performance on all of these tests (O-Dotting, Direction Marking, Tracing and Gate Making) is dependent in large part on the willingness of the subject to stick at a speeded, repetitious and relatively monotonous task.

While the nature of this factor can be no more than speculated on from the results of the present study, it appears most interesting and is important in the Submarine School criterion.

FACTOR V: COMPLEX VISUALIZATION

Direction Making	.77
Thurstone Space	.44
Thurstone Reasoning	.44
Memory for Diagrams	.42
Counting	.37
Line Following	.37
Thurstone Number	.34

Factor V is difficult to interpret. It is probably a composite of three or four factors which are somehow related and all measured in some degree by the

Direction Marking test. Direction Marking or direction plotting has been found to have high loadings on the Visualization factor in a number of studies. Roff[2] found direction marking tests had high loadings on a factor which he called Complex Perception. Attention and memory tests also were loaded on this factor.

Since space, reasoning, memory and number all seem to be involved in Factor V, it must be concluded that the factor is complex. Because of the visualizing nature of the Direction Marking and several of the other tests, the factor is tentatively labeled COMPLEX VISUALIZATION. Its nature should be explored more fully in a more definitive analysis.

FACTOR VI: NUMBER

Final Average	.63
Aptitude for S/M	.56
Thurstone Number	.53
Operating Average	.51
Navy Clerical	.49
Coding	.39
Examination Average	.39
Navy ARI	.32

All of the aptitude tests with appreciable loadings on this factor clearly have number content. The high loadings indicated for the two Submarine School criteria, Final Average and Aptitude for Submarines are interesting and plausible. During the course of training, a submarine candidate is required to learn a great many things for which number ability would be important. Examples of these are learning the frame numbers of the submarine, remembering the cubic footage of certain tanks, the pressures in various lines, being appraised of operating speeds and depths, calculating weights of given volumes of sea water, fresh water, fire, oil, etc. This lends support to the proposed identity of this factor and points up

[2] Roff, M.F., Personnel Selection and Classification Procedures: Perceptual Tests, A Factor Analysis, Project Report, Project No. 21-02-009, USAF School of Aviation Medicine, Randolph Field, Texas.

the importance of Number variance in the Submarine School criterion.

FACTOR VII: VISUAL-REASONING

Navy ARI	.56
Space-O	.48
Thurstone Reasoning	.44
GCT	.36
Memory for Diagrams	.32
Navy Clerical	.32
Navy Mechanical	.24

Factor VII also appears to be complex, clearly having both reasoning and visual components. The entire Navy basic battery is loaded on this factor, but predominantly ARI, together with Thurstone's Reasoning from the Primary Mental Abilities battery and the Spatial Orientation test which is thought to involve both visualization and reasoning to an appreciable extent.

FACTOR VIII: CAREFULNESS? CONSERVATIVENESS?

Navy Clerical	.78
Simple Detail	.60
Social Science Bluffs (lack of bluffing)	.41
Tracing	.35
Line Following (Visual Attention)	.31
Thurstone Verbal	.31
Counting	.29

Factor VIII offers an interesting hypothesis although nothing definite about its nature can be firmly established on the basis of this study. At first glance the factor appears to involve perceptual speed. The high loading of (lack of) Social Science Bluffs, however, which was a score based on items which had no answers (and which therefore, according to directions, should have been omitted by the respondent) demands special interpretation.

The Carefulness factor previously has been defined largely by wrong scores where high scores reflect lack of carefulness. Since wrong scores were not included in this analysis, the similarity between this and other factors described

as Carefulness cannot be determined. The substantial negative correlations between total and wrong scores on such tests as Simple Detail, Counting, and Tracing, plus the high negative relationship between total score and bluffing score on the Social Science test, lend support to the notion that high total scores on these tests may reflect conservativeness or carefulness of response as well as the other abilities associated with them. Further work in an effort to test this hypothesis would seem worthwhile.

Chapter VI

PREDICTION OF PERFORMANCE ABOARD SUBMARINES AS DETERMINED BY RATINGS AND TESTS

Ratings of General Traits

One of the initial hypotheses of this study was that variance of an attitudinal or adjustment nature might be equally or more important for success aboard submarines than that of a technical or job knowledge nature. In order to assess both the technical and the adjustment aspects of performance in an economical fashion, it had been decided that some sort of ratings by superiors must necessarily form at least a part of the criterion measures of shipboard performance. Furthermore it was decided that this rating device should be designed to tap as many important aspects of shipboard performance as was practical and economical.

The shipboard rating scale finally developed was what has been referred to as a man-to-man scale. It was so named because it permitted, in fact required, that all men in a group be rated at the same time on a given trait, in contrast to the majority of rating scales which require that a given man be rated on all traits or all items at some point in time before the next man is rated. This man-to-man format helped, it is thought, the raters compare the various men in their groups and decide whether or not Smith had more of this ability or this trait than Jones and, to some extent, how much more.

Working with Naval personnel to obtain leads as to the important shipboard behaviors to be included in this scale, and later performing a factor analysis of intercorrelations of traits of the earlier forms of the scale, the following ten traits were finally arrived at for inclusion in the final form of the rating scale.

1) Cooperation
2) Knowledge of the Job
3) Discipline
4) Application and Initiative
5) Judgment and Common Sense
6) Leadership
7) Neatness of Work
8) Care of Equipment
9) Ability to Troubleshoot
10) Sincerity in Doing the Job

The factorial analysis of the predecessors of this final scale had indicated, on two separate occasions, that at least two broad general areas of performance had been rated. One of these was characterized by high scores in such traits as Knowledge of the Job, Leadership, Ability to Troubleshoot and the ratee's pay grade. This factor had been regarded as primarily technical in nature and labeled as Technical Competence. The two traits having highest loadings on this factor were Knowledge of the Job and Ability to Troubleshoot.

A second factor which had emerged in the two independent analyses was characterized by such traits as Cooperation, Application and Initiative, Discipline, Care of Equipment and Sincerity in Doing a Good Job. This factor clearly seemed to reflect more in the way of adjustment or attitude toward submarine life than did the previous one. It was, therefore, given the tentative label Personal Adjustment.

The reliability of the ratings obtained with this form was generally found to be satisfactory both from the standpoint of inter-rater agreement and from the standpoint of rate-rerate reliability. Inter-rater agreement ranged from about .45 to .65 for individual traits and better than .70 for total score. Rate-rerate reliability measured over an elapsed period of five to nine months was above .80.

Ratings were accomplished by three of each man's supervisors. For the most part, these were the man's division officer and the two leading petty officers, usually chiefs of his gang.

At this point, it is necessary to interject some information about the nature of the men who made up the subjects for the remainder of the study. The two largest rates aboard submarines of the Navy, in terms of the number of men aboard, were the Electrician's Mates and the Enginemen. Since it would have been impractical to develop separate criterion measures for all of the various rates aboard, it was decided to concentrate on these two rates because of their relatively large numbers. Furthermore, since our interest was primarily with those men to whom aptitude tests had been administered at New London, we were confined not only to

the two rates of Electrician's Mates and Enginemen but also to relatively recent graduates of the Submarine School.

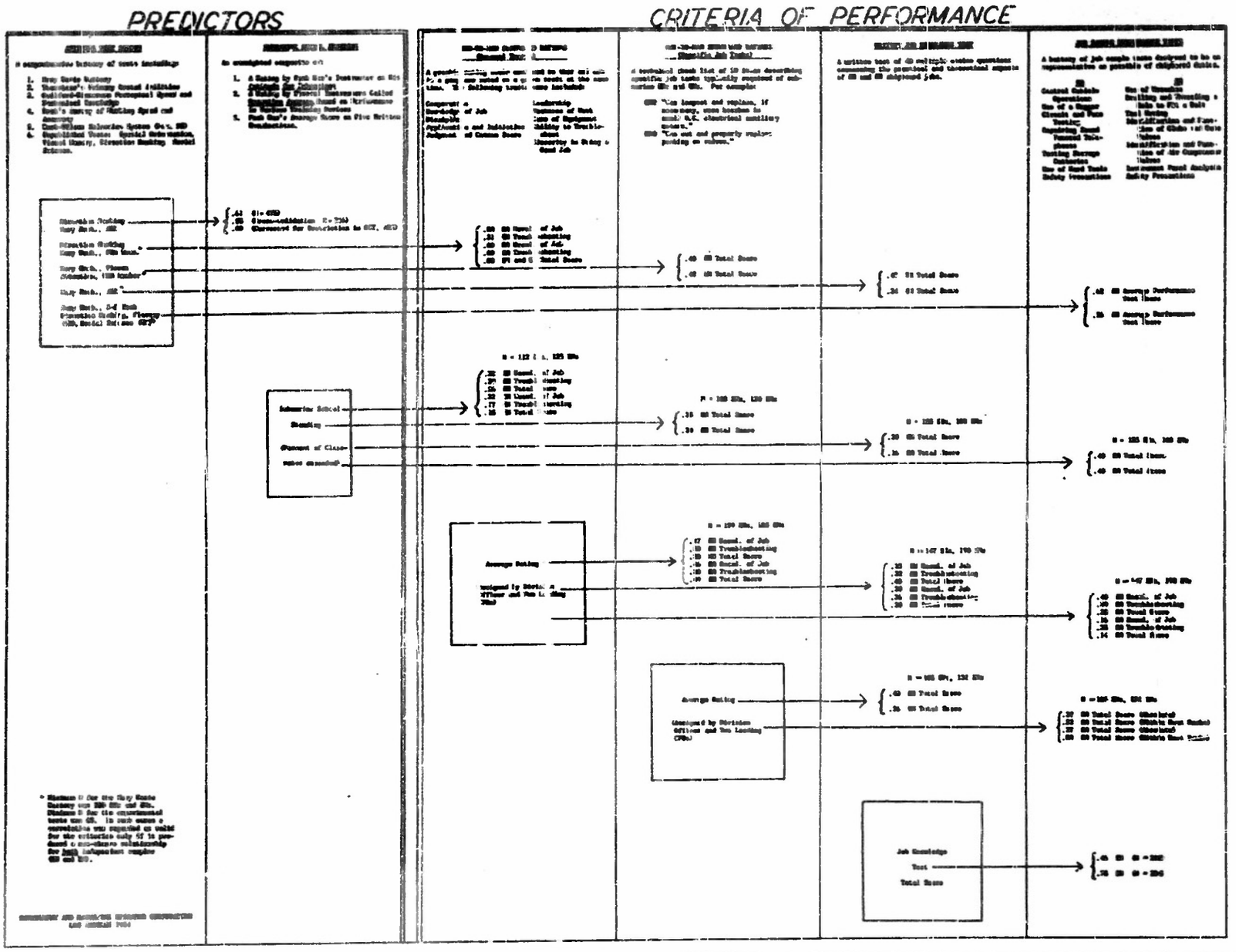
It was arbitrarily decided that no man would be included in the study who had not served aboard a boat at least six months. It was presumed that his supervisors would have had a reasonable chance to become acquainted with him in such a period of time. The gathering of criterion measures was accomplished mainly in the summer of 1951. This meant that the subjects for study were, with few exceptions, Strikers and 3rd Class Petty Officers. They represented men with from six months' to approximately two years' experience aboard submarines. This selection of subjects no doubt placed certain restrictions on the ranges of abilities represented and on the extent to which each man might have had an opportunity to acquaint his superiors with his capabilities.

In the chart on page 42, the interrelationships between all predictors and criteria in this study are presented. In the first column, under Criteria of Performance, the first group of multiple R's reveals the extent to which ratings on the general traits of the rating scale were predictable from aptitude test scores.

Most revealing in this set of relationships, perhaps, is the lack of predictability of total score on the rating scale. None of the aptitude tests showed more than chance deviations from zero correlations for either the Electrician's Mates or Enginemen population. Since the predictors were aptitude test scores, it was thought that scores on the technical traits of the rating scale might be predictable whereas scores on personal adjustment traits might not. To a small extent, this proved true. The Direction Marking, Naval Mechanical and PMA Reasoning test, in combination, produced multiple R's of .28 and .31 for the Electrician's Mates ratings on Knowledge of the Job and Troubleshooting. This relationship did

PREDICTION AND MEASUREMENT OF PERFORMANCE

OF SUBMARINE ELECTRICIAN'S MATES AND ENGINEMEN



not hold up for the Enginemen, however, which must be regarded as something of a failure of cross-validation.

Ratings on Specific Job Tasks

In order to get as close as possible to specific shipboard behaviors which should be observable to supervising officers and petty officers, a second rating scale had been designed which also utilized the man-to-man format. This consisted of items of specific technical behavior in contrast to the general traits contained in the rating scale previously described. One of these specific check list rating scales was developed for the Electrician's Mates and one for the Enginemen. The items selected were those which, according to shipboard personnel, were of an appropriate difficulty level for Strikers and 3rd Class Petty Officers, were frequently performed aboard ship, and which raters would know whether or not his men could perform. For example, the Electrician's Mates scale included items such as: "Can determine if a submarine battery needs water", "Can replace a length of ship's lighting cable", "Can rewire electric switches", "Can parallel D.C. generators using correct procedure", "Can stand a repair watch in the battery compartment during battle stations", etc. Typical items from the Enginemen's check list were: "Can stand a throttle watch on the main engines", "Can cut and properly replace packing on valves", "Can line up and start a main engine", "Can operate vapor compression units", "Can make necessary adjustments or replacements in main hydraulic systems", etc.

Fifty such items were selected after pre-testing for both the Electrician's Mates and Enginemen. Again, the division officer (Engineering Officer) and the two leading Chief Petty Officers of each man's gang performed the rating. Inter-rater agreement on total scores for the Check List rating scale proved to be above .80. Some officer ratings had to be discarded because the officers found that they did not have sufficient specific knowledge to utilize the check list.

As can be seen in the second column of the chart under Criteria of Performance, the specific check list ratings proved somewhat more predictable by aptitude tests than did the ratings on more general traits of behavior. The Navy Mechanical test, the Visual Attention test, and PMA Numerical test combined to yield a multiple R for both independent samples in the neighborhood of .40.

Written Job Knowledge Test and Job Sample Performance Tests

While ratings were easily obtainable and, it was hoped, would eventually be shown to be suitable criteria of shipboard performance, the many questions usually raised about their usefulness led to the development of actual tests of performance which could be administered to the subjects aboard ship. As can be seen in the last two columns of the chart under Criteria of Performance, these consisted of a Written Job Knowledge test of a conventional multiple-choice variety and actual Job Sample tests which required each man to demonstrate his skill at selected tasks representative of what the Electrician's Mate or Engineman actually does aboard ship.

The Electrician's Mate was tested on Control Cubicle Operations, Use of a Megger, Circuit and Fuse Testing, Repairing Sound Powered Telephones, Testing Storage Batteries, Use of Hand Tools, and Safety Precautions in connection with his work. The Engineman was tested on Use of Wrenches, Drilling and Threading a Hole to Fit a Bolt, Naming of Tools associated with his work, Identification and Function of Globe and Gate Valves, Identification and Function of Air Compressor Valves, Instrument Panel Analysis, and Safety Precautions associated with his work. The performance test battery required approximately two hours' administration time. The reliabilities of the sub-tests ranged from .50 to .80 and the reliability of the composite battery for EM's and EN's was estimated at about .90.

The Written Job Knowledge test contained multiple-choice questions concerning the practical and theoretical aspects of the Electrician's Mate and Engineman job.

The reliabilities of these tests were .86 and .68 respectively. As can be seen in the chart, the aptitude tests predicted performance on these tests to a reasonable degree. The Navy Mechanical and Navy Arithmetic tests combined to give multiple R's of .47 and .34 for the Electrician's Mates total score and Enginemen total score respectively.

More interesting, perhaps, is the extent to which the Navy Mechanical, Guilford-Zimmerman Mechanical and Direction Marking tests combined with Verbal Fluency (for the Electrician's Mates) and Social Science (for the Enginemen) to yield multiple R's of .62 and .56 for the Electrician's Mates and Enginemen's average performance test scores respectively. No particular hypothesis is advanced to explain the presence of the two verbal tests of these batteries. Their contribution to the total was relatively small. However, the presence of the Navy Mechanical and the Direction Marking tests which had proved predictive of both general and specific ratings, and Written Job Knowledge test scores as well, established some confidence that tests of this type are valid for many of the technical aspects of shipboard performance aboard submarines.

Submarine School Standing as a Predictor

Submarine School Standing appears to be a reasonable predictor of subsequent performance aboard submarines. While the relationships shown in the second row of the chart are not particularly high, it must be remembered that we had dealt with a relatively restricted range of performance in confining the study to 3rd Class Petty Officers and Strikers. It is considered quite possible that, had the careers of these men been followed until they became 2nd Class or 1st Class Petty Officers, the relationship of Submarine School Standing to subsequent performance might have proved higher.

In observing the relationships between Submarine School Standing and the criteria developed for this study, a trend can be observed that was generally true

of all the relationships determined. This is the fact that relationships between predictors and criteria were generally higher for the Electrician's Mates than for the Enginemen. Submarine School Standing predicted Electrician's Mates Job Knowledge to the extent of .30, Troubleshooting .29, and Total Score .26. In contrast, it predicted Enginemen Job Knowledge only .21, Troubleshooting .17, and Total Score .15.

The same trend can be observed in the correlation between Submarine School Standing and specific check list rating (.35) for Electrician's Mates and only .20 for Enginemen. Again, the trend can be observed with the Written Job Knowledge test, the correlation being .33 for Electrician's Mates and .16 for Enginemen. When it came to total score on the Job Sample performance tests, however, Submarine School Standing proved to be equally valid for the two rates, the correlation being .40 in each case.

This general trend toward greater validity for Electrician's Mates ratings is believed to be due to the fact that the Electrician's Mate frequently works on isolated, rather specific, and comparatively easily observed tasks. He much more frequently works alone than does the Engineman. He usually works in a fairly confined area or on a relatively small piece of equipment. The Engineman in contrast, particularly at the Striker and 3rd Class levels, tends to work with others on highly complex and large pieces of machinery. It is quite possible that the extent of his contribution to the total performance is relatively difficult to evaluate because he so often performs as a member of a team with a great deal of help and direct supervision from others.

Relationship of Ratings on General Traits to Other Criterion Measures

In the third row of the chart, the extent of relationships between ratings on the ten general traits and the other criterion measures can be inspected. It will be noticed first of all that there were moderately high relationships between ratings on the general trait rating scale and ratings on the specific check list

rating scale. Total scores on these measures correlated .58 for the Electrician's Mates and .49 for the Enginemen. Scores on the general traits, Knowledge of the Job and Troubleshooting, correlated .47 and .43 with the total check list scores for the Electrician's Mates and .46 and .32 for the Enginemen. It was generally observed throughout the study, as might be expected, that the two different kinds of ratings correlated higher with each other than either did with any other measure.

It will be noticed in the next column that ratings on Knowledge of the Job and Troubleshooting correlated to some extent with Written Job Knowledge test scores. These correlations were .25 and .32 for the Electrician's Mates, .32 and .36 for the Enginemen. The correlation of Written Job Knowledge test scores with Total Score on the general rating scale was considerably lower, dropping all the way to .05 for the Electrician's Mates and .20 for the Enginemen.

Finally, the relationship of scores in the general rating scale traits to the Job Sample test scores can be seen in the last column. The picture here is encouraging in the case of the Electrician's Mate but, unfortunately, discouraging in the case of the Engineman. Ratings on the traits Knowledge of the Job and Troubleshooting correlated .48 and .39 with total Job Sample test scores for the Electrician's Mates. Considering the restriction in range imposed by the pay grades of the men in the study, this is a fairly substantial result. For the Enginemen, however, the correlations for the same two traits dropped to .16 and .23 respectively. Again, the hypothesis is advanced that the general nature of the Engineman job, at the level studied, made it rather difficult for raters to know just how much in the way of technical ability the Enginemen really possessed.

Relationship of Specific Check List Ratings to Other Criterion Measures

In the fourth row of the chart, it will be observed that the more specific and technical check list ratings did not fare very much better in their prediction of performance on Job Sample and Written Job Knowledge tests than did the technical traits of the general rating scale. Total scores on the check list rating scale

correlated .40 with total scores on the Written Job Knowledge tests for Electrician's Mates but only .26 for the Enginemen.

Since the check list ratings represented the raters' opinions of the ability of their men to perform important job tasks, a somewhat exhaustive study was performed of the relationship between these ratings and performance as revealed by the Job Sample tests. First, check list scores were taken as absolute values and correlated with Job Sample performance test scores. This procedure resulted in a correlation between the two measures of .37 for Electrician's Mates and .27 for the Enginemen.

This somewhat surprisingly low relationship was thought possibly to be due in part to differences in the leniency of raters in the various submarines that were included in the sample. Therefore, the rank order of each man within his own boat on the performance test was correlated with his rank order on the check list ratings. The average rank order coefficient obtained over about a dozen submarines was .52 for the Electrician's Mates and .30 for the Enginemen. In the case of the Electrician's Mates at least, this procedure seemed to increase the validity of the check list ratings to a point where they might be considered to yield a reasonable indication of technical skill where Job Sample test scores are considered a good criterion of those skills.

In order to shed further light on the validity of raters' opinions about the skills of their men, items from the check list ratings which were identical with, or similar to, sub-tests in the performance test battery were correlated with scores on their counterparts in the performance test. The median correlation between rated ability to do a task and tested ability to do the same or a very similar task was .18 for both Enginemen and Electrician's Mates. This strongly suggests that most raters in this study, in spite of the fact that they lived with their men 24 hours a day, did not have very much knowledge of whether or not their men could actually perform specific tasks which were regarded as important parts of the Electrician's

Mate and Engineman jobs. While total scores on the check list did have some, and perhaps a usable, validity in predicting total scores on Job Sample tests, the statement on the part of a rater that his men can or cannot do a specific task must be regarded with some suspicion.

Relationships Between Written Job Knowledge Test Scores and Job Sample Performance Test Scores

Finally, in the last row of the chart can be observed the extent to which scores on the multiple-choice Written Job Knowledge test correlated with total scores on Job Sample tests. The magnitude of the correlations obtained, .46 for Electrician's Mates and .35 for Enginemen, would hardly lead one to accept Written Job Knowledge test results as a substitute for whatever is measured by actual Job Sample tests.

This general magnitude of correlation, in the neighborhood of .40, was about the same as the extent of the correlation observed among the various parts of the performance test battery itself. This raises some interesting questions as to how many Job Sample tests one would have to assemble to cover completely the technical job skills which Navy men of a given rate are supposed to possess. It raises additional questions about the amount of specific variance associated with the Job Sample sub-tests and whether or not there are general skills which might be measured which would be independent of the specific materials or equipment associated with a given task. It also raises questions concerning the extent to which training and subsequent specialization aboard the submarine will limit the relationship between Job Sample tests of all types.

Conclusions

The following conclusions are drawn from the results of this study:

1. Scores from a battery of aptitude tests, including mechanical, arithmetic, and direction marking type tests, will predict to a large extent the reliable variance in the Submarine School criterion.
2. Scores from a battery of aptitude tests, including mechanical, arithmetic, reasoning, visual attention, direction marking and possibly verbal tests, will predict to a moderate extent the technical aspects of shipboard performance as measured by job sample tests, written job knowledge tests, and specific check list ratings.
3. A large part of the variance in shipboard ratings of performance appears to be non-technical in nature, suggesting that ratings measure to some extent, attitude toward the job or adjustment to Navy life.
4. Aptitude test scores, of the type employed in this study, apparently will not predict the non-technical aspects of performance measured by the ratings.
5. Submarine School standing appears to be a fair predictor of subsequent shipboard performance as measured by job sample tests and ratings.
6. In this study the performance of Electrician's Mates was more reliably measured and better predicted than that of Enginemen. This suggests that the factor of rate may play a major role in determining the type of criterion measure that can be used and the extent to which it will be predictable from other measures. (For example, the fact that Enginemen tend to work in small groups and

on large pieces of equipment, in contrast to Electrician's Mates who frequently work alone and on smaller pieces of equipment, is believed to make it more difficult for an observer to judge the skill and knowledge of a particular Engineman.)

7. Scores from written job knowledge tests, as developed for this study, do not correlate high enough with job sample test scores that written test results can be accepted as a satisfactory indication of a man's ability to perform the practical factors of his rate.

APPENDIX

NEW LONDON INTERCORRELATION MATRIX
(Tetrachoric Coefficients)

N = 475 For Variables 1-20, 24-27
N = 319 For Variables 21-23

1	OCT	-													
2	NAVY MECH	40													
3	NAVY ARI	49	15												
4	NAVY CLER	01	-02	23											
5	PIA VERBAL	47	00	20	35										
6	PIA REASON	37	30	42	36	34									
7	PIA SPACE	14	44	07	14	25	46								
8	PIA NUMBER	14	-19	26	45	34	33	11							
9	PIA FLUENCY	25	-03	22	26	22	26	08	22						
10	GATE MAKING	15	-10	00	23	25	23	16	20	22					
11	LINE FOLLOWING	18	23	02	-15	-09	09	09	-09	00	-05				
12	SIMPLE DETAIL	-09	04	-15	39	17	23	40	22	13	25	-07			
13	TRACING	-05	11	-07	20	-09	06	13	03	12	31	08	29		
14	CODING	08	-13	11	42	24	29	27	41	13	21	-04	23	07	
15	O - DOTTING	-01	00	-10	19	13	15	20	09	15	42	-15	23	35	20
16	COUNTING	12	05	17	49	21	43	32	36	14	30	04	39	20	47	34
17	G-Z PERC. SPEED	13	19	20	38	14	36	30	16	10	35	07	35	11	40	18
18	G-Z MECH KNOWL.	33	72	07	-22	-03	07	17	-19	06	01	16	-07	03	-03	-05
19	SPATIAL ORIENTATION	42	44	25	00	13	40	29	-11	13	10	28	04	09	-06	13
20	MEMORY FOR DIAGRAMS	10	07	22	17	12	29	15	06	00	17	09	11	03	16	05
21	DIRECTION MARKING	31	19	21	-08	14	39	37	43	20	33	07	06	29	27	-03
22	SOC. SCIENCE	72	07	24	-19	43	41	20	16	30	21	13	-07	00	19	-11
23	SOC. SCIENCE BLUFFS	-34	03	-25	-61	-57	-32	-06	-30	-31	04	11	-18	-07	-22	-22
24	APT. FOR S/M	36	40	25	00	18	19	29	17	01	14	07	09	-03	18	04
25	OPER. AVERAGE	30	36	34	19	06	16	15	05	04	06	12	03	03	10	02
26	EXAM. AVERAGE	49	38	36	15	27	27	21	09	11	17	08	08	04	11	-04
27	FINAL AVERAGE	39	43	34	09	13	23	26	13	05	14	09	05	09	13	03
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

NEW LONDON INTERCORRELATION MATRIX (cont.)

43											
00	17										
07	31	23									
25	19	- .	30								
35	24	21	23	40							
16	16	09	23	36	21						
-24	-16	00	06	06	-26	-68					
04	15	29	29	06	29	27	00				
08	19	30	26	08	16	16	-14	79			
14	14	33	31	07	34	19	-11	64	63		
13	14	32	18	03	46	18	-05	95	91	93	-
16	17	18	19	20	21	22	23	24	25	26	27

NEW LONDON BATTERY
CENTROID FACTOR MATRIX

TEST		FACTOR I	II	III	IV	V	VI	VII	VIII	h^2
GCT	1	59	-41	-43	-24	-03	-22	08	-05	.817
NAVY MECH	2	42	-48	30	-37	-41	16	10	12	.852
NAVY ARI	3	44	-20	-35	-13	19	16	-26	08	.509
NAVY CLER	4	46	50	-19	39	-31	32	-37	23	1.038
PIA VERBAL	5	47	15	-42	07	-13	-11	22	10	.512
PIA REASON	6	65	13	-14	-29	07	17	-10	15	.609
PIA SPACE	7	49	08	21	-18	-07	19	23	07	.422
PIA NUMBER	8	39	38	-25	19	27	15	20	11	.543
PIA FLUENCY	9	32	12	-25	04	-08	-18	-07	-03	.226
CATE MAKING	10	37	26	14	09	07	-26	-16	-27	.403
LINE FOLLOWING	11	43	17	24	-23	-16	05	09	07	.365
SIMPLE DETAIL	12	35	47	28	10	-31	12	05	06	.548
TRACING	13	24	25	35	13	-13	-28	-10	09	.373
CODING	14	43	38	-05	09	17	19	10	-24	.472
O - DOTTING	15	32	12	23	15	13	-32	-14	04	.333
COUNTING	16	54	46	10	-03	09	18	-02	-06	.559
G-Z PERC. SPEED	17	51	20	13	-12	08	19	-15	-45	.459
G-Z MECH KNOWL.	18	31	-45	24	-26	-26	-06	17	-21	.568
SPATIAL ORIENTATION	19	43	-22	15	-33	-11	-09	-30	08	.481
MEMORY FOR DIAGRAMS	20	33	08	05	-31	31	-06	-19	13	.367
DIRECTION MARKING	21	60	15	26	-13	33	-29	27	30	.823
SOC. SCIENCE	22	53	-07	-45	-24	14	-40	17	-17	.783
SOC. SCIENCE BLUFFS	23	43	17	-64	10	-27	-09	20	10	.764
APT. FOR S/I	24	56	-56	12	34	18	19	22	-08	.880
OPER. AVERAGE	25	54	-54	11	37	08	20	-09	-06	.790
EXAM. AVERAGE	26	64	-48	08	30	04	-04	-03	19	.777
FINAL AVERAGE	27	64	-61	20	44	28	15	13	13	1.150

NEW LONDON BATTERY
ROTATED ORTHOGONAL FACTOR MATRIX

TEST		FACTOR I_{10}	II_{11}	III_{12}	IV_{11}	V_{11}	VI_{9}	VII_{10}	$VIII_{8}$	h^2
GCT	1	43	04	70	09	12	10	33	-10	.828
NAVY MECH	2	83	-02	-03	-08	22	-18	24	17	.863
NAVY ARI	3	12	06	25	03	11	32	56	-09	.516
NAVY CLER	4	-19	20	-03	05	-07	49	32	78	1.037
PMA VERBAL	5	04	04	56	00	15	27	02	31	.509
PMA REASON	6	13	27	22	-05	44	23	44	20	.621
PMA SPACE	7	34	22	01	-07	44	12	01	23	.431
PMA NUMBER	8	-23	15	22	00	34	53	-05	17	.553
PMA FLUENCY	9	-03	15	36	13	01	09	13	19	.231
GATE MAKING	10	-06	50	11	35	06	00	00	11	.404
LINE FOLLOWING	11	24	27	01	-02	37	-05	06	31	.370
SMALL DETAIL	12	04	32	-13	02	20	07	-11	60	.536
TRACING	13	-01	20	-07	39	14	-15	-07	35	.367
CODING	14	-08	49	10	-04	21	39	-06	12	.472
8 - DOTTING	15	-03	20	03	31	18	-01	02	12	.349
COUNTING	16	-05	51	01	02	37	27	08	29	.563
G-Z PERC. SPEED	17	19	56	-01	-02	13	13	20	21	.468
G-Z MECH KNOWL.	18	70	14	11	02	06	-20	01	-06	.570
SPATIAL ORIENTATION	19	36	16	06	18	19	-19	48	08	.500
MEMORY FOR DIAGRAMS	20	-05	19	06	14	42	-02	36	-09	.376
DIRECTION MARKING	21	10	13	19	40	77	06	-05	09	.832
SOC. SCIENCE	22	11	23	79	12	23	04	13	-14	.794
SOC. SCIENCE BLUFFS	23	-01	-02	69	-11	01	31	06	41	.757
APT. FOR S/H	24	65	02	04	32	10	56	-03	-18	.883
OPER. AVERAGE	25	58	05	-04	39	-07	51	19	-08	.800
EXAM. AVERAGE	26	53	-08	13	51	12	39	20	07	.776
FINAL AVERAGE	27	64	-10	-03	53	20	63	05	-15	1.163