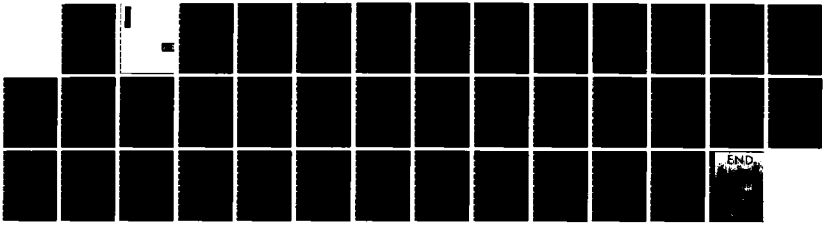
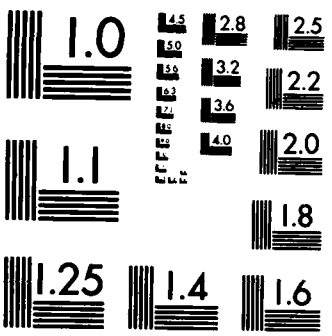
AD-A136 976 EVALUATION OF PREDICTORS AND CRITERIA FOR JOB PERFORMANCE OF GENERAL DETA..(U) NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN DIEGO CA C H CORY

1/1

UNCLASSIFIED NOV 83 NPRDC-TR-84-8 F/G 5/9 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NPRDC TR 84-8

November 1983

EVALUATION OF PREDICTORS AND CRITERIA FOR JOB PERFORMANCE OF GENERAL DETAIL (GENDET) PERSONNEL

Charles H. Cory

Reviewed by
Martin F. Wiskoff

Released by
J. W. Renard
Captain, U.S. Navy
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NPRDC TR 84-8	2. GOVT ACCESSION NO.: AD-A136 976
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): EVALUATION OF PREDICTORS AND CRITERIA FOR JOB PERFORMANCE OF GENERAL DETAIL (GENDET) PERSONNEL	5. TYPE OF REPORT & PERIOD COVERED: Technical Report FY79-81
	6. PERFORMING ORG. REPORT NUMBER: 62-83-02
7. AUTHOR(s): Charles H. Cory	8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Navy Personnel Research and Development Center San Diego, California 92152	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: ZF63-521-030-01.01
11. CONTROLLING OFFICE NAME AND ADDRESS: Navy Personnel Research and Development Center San Diego, California 92152	12. REPORT DATE: November 1983 13. NUMBER OF PAGES: 38
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

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

Job performance
Surrogate criteria
Test validation
Personnel selection
Supervisory evaluations
Cross validation
Predictive validity

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

This research evaluated written aptitude tests and biographical questionnaires for selecting and assigning personnel to general detail (GENDET) billets in the Navy. Records of 46,231 GENDETs were extracted from 10 data bases used for previous research. Personnel were classified into six mental ability levels, and statistics describing the job outcome characteristics of the groups were computed. Predictive validity analyses were conducted. Results showed that mental ability had a moderately high relationship with achievement of rated (E-4 and above) status, highest pay grade

achieved, and days required to achieve rated status; and a low relationship to attrition and overall behavior, as measured by disciplinary infractions. The predictability of job outcome variables was generally much higher and somewhat more stable than was the predictability of supervisors' ratings of job performance.

FOREWORD

This research, which was conducted in support of project ZF63-521-030-01.01 (Prediction of Performance), was designed to improve the use of personnel who are assigned to general detail (GENDET) billets in the fleet. The results are intended for use by enlisted personnel detailers, fleet personnel concerned with assignment of GENDETs, personnel researchers, and cognizant officers in the military personnel and Navy recruiting commands.

Appreciation is extended to Dr. Eric (E. K.) Gunderson and Mr. Mylan Miller, both of the Naval Health Research Center, for providing enlisted history data for personnel in the data base.

J. W. RENARD
Captain, U.S. Navy
Commanding Officer

JAMES W. TWEEDDALE
Technical Director

SUMMARY

Problem

General detail (GENDET) recruits, who do not receive technical school ("A" school) training after recruit training, characteristically have higher attrition, slower advancement, and lower first-term reenlistment rates than do recruits who receive technical training. Contributing factors seem to include the following: (1) GENDETs' mental abilities, on the average, are lower than those of personnel with "A" school training, and (2) assignments for GENDETs are not so carefully matched to their individual capabilities as are those of personnel sent to "A" school. While "A" school assignments are based on scores drawn from scientifically derived prediction equations, GENDET billet assignments result from informal matching of persons to jobs. To increase the usefulness and retention of GENDETs, improved assignment and classification measures are needed.

Objective

The objective of this research was to improve classification and assignment of GENDETs by (1) identifying the most appropriate criteria for evaluating GENDET performance, (2) identifying GENDET performance and advancement characteristics, (3) determining whether GENDETs could be selected more accurately by using separate predictor composites for seaman (SN), fireman (FN), and airman (AN) apprenticeship ratings than by using a single predictor composite for the whole group, and (4) evaluating selection measures and predictor composites that could be used for assigning personnel to GENDET billets.

Approach

Records of 46,231 personnel assigned as GENDETs were extracted from 10 data bases that had been assembled from previous research efforts at the Navy Personnel Research and Development Center. From this total, the 39,019 personnel who had Armed Forces Qualification Test (AFQT) scores were classified into six mental level groups (1s, 2s, High 3s, Low 3s, High 4s, and Low 4s); statistics describing the rating progression and job performance characteristics of the groups were computed. Predictive validity analyses were carried out on the personnel in the eight largest data bases (sample sizes ranged from 1,243 to 15,041), using a double cross-validation paradigm.

Findings

1. Mental ability as measured by AFQT scores had a moderately high association with achievement of rated (E-4 and above) status (BIRTD), highest pay grade achieved (HIPG), and days required to achieve rated status (DAE4). Mental ability had a low relationship to attrition (ATTR) and to overall behavior as measured by disciplinary infractions (BEHR).

2. On the average, the percentage of a 4-year enlistment during which personnel performed at the E-3 level or higher varied little among the AFQT mental level groups. In contrast, the average percentage of a 4-year enlistment during which personnel performed at the E-4 level and above ranged monotonically from 46 percent for Mental Group 1 to 0 for Low 4. On the average, Mental Group 1 and 2 personnel spent 9 and 3 percent respectively of the 4-year enlistment at the E-5 level.

3. Non-school-eligible (NSE) personnel (AFQT mental level groups Low 3 and 4) constituted 28 percent of those who achieved a pay grade of E-4 and above, appearing in all 58 ratings studied. Nine ratings had greater concentrations of NSE than school-eligible personnel. Most rated NSE personnel were in mess management specialist, hospital corpsman, and dental technician ratings.

4. Although a comprehensive series of analyses was carried out to develop separate predictive composites for the SN, FN, and AN apprenticeship ratings, these predictive composites were no more accurate on cross-validation than was a single composite computed for the combined group. Thus, it was not possible to improve predictive accuracy by developing separate predictor composites for each apprenticeship rating.

5. The predictability of job performance criteria (HIPG, BIRTD, BEHR, and DAE4) derived from Navy records was generally much higher and was somewhat more stable than was the predictability of supervisors' ratings of job performance (OVER).

6. OVER had a low relationship to such outcome criteria as HIPG and BIRTD. In fact, OVER did not predict career outcome criteria as well as did AFQT.

7. Although 125 experimental predictors were evaluated, nearly all of the predictiveness of the battery was supplied by variables that were available operationally; that is, scores from either the Armed Services Vocational Aptitude (ASVAB) or its Navy predecessor, the Basic Test Battery, scores from success chances of recruits entering the Navy (SCREEN), or biographical variables such as age and years of education.

Conclusion

The findings of this research indicate that, so long as there is an excess of applicants to GENDET billets, the quality and subsequent job performance of GENDET personnel can be improved by using predictor composites to select them.

Recommendations

1. Future validations against GENDET job performance should use DAE4 and either HIPG or BIRTD as criteria instead of OVER.

2. In future research using job performance criteria to select GENDETs, a single predictor composite for the combined SN, FN, and AN ratings should be computed.

3. To maximize the percentage of GENDETs who achieve rated status, selection should be based on ASVAB and SCREEN scores. To maximize the behavioral characteristics of GENDETs, selection should be based on SCREEN score and age.

CONTENTS

LIST OF TABLES

INTRODUCTION

Problem

Personnel who are not assigned to technical training in "A" schools directly after they complete boot camp are assigned to seaman (SN), fireman (FN), or airman (AN) apprenticeship school for orientation and training in the basic skills needed to function as general detail (GENDET) personnel. After they complete apprenticeship training, these personnel are assigned to the fleet in GENDET positions, where they receive on-the-job training for specific ratings and perform necessary housekeeping chores. About 30 percent of personnel who enlist in the Navy are assigned as GENDETs (Cory, 1982).

GENDETS characteristicallv have higher attrition, slower advancement, and lower first-term reenlistment rates than do recruits who receive training in "A" school. Contributing factors seem to include the following:

1. The mental abilities of GENDETs, on the average, are lower than those of personnel with "A" school training.

2. Assignments for GENDETs are not so carefully matched to their individual capabilities as are those of personnel sent to "A" school. While "A" school assignments are based on scores drawn from scientifically derived prediction equations, GENDET billet assignments result from informal matching of persons to jobs.

To increase the usefulness and retention of GENDETs, improved assignment and classification measures are needed.

Background

The Navy Personnel Research and Development Center (NAVPERSRANDCEN) began research on selection of GENDETs in 1967 in connection with Project 100,000, which concerned job performance characteristics of personnel whose Armed Forces Qualification Test (AFQT) scores[1] fell into Mental Group 4, the 10th through 30th percentiles (Cory, Neffson, & Rimland, 1980). Nearly all Mental Group 4 personnel are assigned to GENDET billets. Cory et al. (1980) investigated whether written test scores were predictive of supervisors' ratings of job performance. They found that the Basic Test Battery (BTB) and Special Classification Tests (SCT), which were used for Navy enlisted classification before the ASVAB was implemented in 1976, were moderately predictive of supervisors' marks or ratings, and that experimental tests did not substantially increase the predictiveness of the operational tests for Mental Group 4 personnel.

Research at NAVPERSRANDCEN that specifically focused on improving the classification and assignment of GENDETs was begun in 1975. The major efforts in this area (Cory, 1982; Siegel, 1983; Siegel & Wiesen, 1977) investigated the use of hands-on performance tests administered in assessment centers as selectors for job assignment purposes. The tentative findings of these studies suggested that performance tests used in conjunction with ASVAB scores can improve selection of GENDETs over ASVAB scores used alone.

---

[1] AFQT scores are now computed from a composite of Armed Services Vocational Aptitude Battery (ASVAB) subtests. However, until January 1976, AFQT was a separate test administered servicewide for military selection.

In conducting GENDET research, it is essential to determine the most appropriate criteria for evaluating job performance. Overall supervisory evaluation (OVER), the most widely used measure and the one that directly evaluates job performance, is plagued with weaknesses associated with the limitations of supervisors as evaluators. Supervisors' marks frequently have been characterized by low interrater reliability, bias, halo effects, and low stability over time (Cory, 1982; Vineberg & Joyner, 1982).

Although indices of enlisted career characteristics appearing in personnel records are only indirect measures of performance, they may be more reliable than are supervisors' marks (Hiatt & Sims, 1980; Lurie, 1981; Vineberg & Joyner, 1982). Many of these indices, such as highest pay grade achieved (HIPG), achievement of rated status (BIRTD), and rate of advancement (DAE4), intuitively appear to be important measures of the value of personnel to the Navy. Thus, they may be better than OVER for use in validating personnel selection tests.

Objective

The objective of the research reported here was to improve GENDET classification and assignment by (1) identifying the most appropriate criteria for evaluating GENDET performance, (2) identifying GENDET performance and advancement characteristics, (3) determining whether GENDETs could be selected more accurately by using separate predictor composites for SN, FN, and AN apprenticeship ratings than by using a single predictor composite for the whole group, and (4) evaluating selection measures and predictor composites that could be used for assigning personnel to GENDET billets.

APPROACH

The research plan was to (1) identify Navy personnel who had been assigned as GENDETs, (2) extract data concerning their job performance characteristics from Navy records, (3) use these data as criteria for regression analyses in developing composites of variables that would be maximally predictive, and (4) cross-validate the predictive composites on holdout samples.

Data Bases

Subjects were selected from 10 data bases assembled in research conducted at NAVPERSRANDCEN during the past 13 years (Cory, 1976; Cory, Neffson, & Dean, 1979; Cory et al., 1980; Dann & Abrahams, 1973; Duffy, Carter, Fletcher, & Aiken, 1975; Siegel & Wiesen, 1977) (see Table 1). Each study involved (1) using experimental predictors in a predictive validity design, (2) collecting criterion data, usually supervisors' marks, 12 to 24 months after enlisted accession, and (3) correlating the predictor and criterion variables.

During the formation of the data bases, it was discovered that, in some instances, personnel in the Project 100,000, Phase 2 study (Cory et al., 1980) had been included in one other study, either Phase 3, 4, or 5 or Project 100,000 or in the Navy vocational interest inventory (NVII) study (Dann & Abrahams, 1973). Thus, to restrict the predictive validity analyses to personnel present in only one data base, 2,216 subjects were dropped from these analyses. However, these personnel were included in the descriptive analyses of GENDET job performance and advancement.

Table 1

Data Bases Used for the Study

Data Base	Year Data Collected	Research Description	Subjects Classified As GENDETs (N)	GENDETs With AFQT Scores (N)	Research Reference
Project 100,000, Phase 1 (P1)	1967	Experimental tests for project.	2,898	2,566	Cory, Neffson, and Rimland, 1980
Project 100,000, Phase 2 (P2)	1968	Experimental tests for project.	6,245	4,850	Cory et al., 1980
Project 100,000, Phase 3 (P3)	1968	Experimental tests for project.	5,992	4,588	Cory et al., 1980
Project 100,000 Phase 4 (P4)	1969	Experimental tests for project.	6,120	4,521	Cory et al., 1980
Project 100,000 Phase 5 (P5)	1970	Experimental tests for project.	3,687	3,287	Not formally reported
Navy Vocational Interest Inventory (NVII)	1969 1970	NVII as predictors of "A" School grades.	17,109	15,041	Dann and Abrahams, 1973
Computerized Perceptual Tests[a]	1972	Computerized tests as predictors of supervisors' marks.	83	83	Cory, 1976
Gates-MacGinitie Reading Test (GM)	1974 1975	Study of reading level of Navy enlistees.	2,761	2,754	Duffy, Carter, Fletcher, and Aiken, 1975
Assessment center study[a]	1975	Hands-on performance tests for classifying GENDETs.	86	86	Siegel and Wiesen, 1977
Boeing Computer Services/ Cleff Inventory (BCS/ Cleff)	1976	Interest inventory for "A" School selection	1,250	1,243	Cory, Neffson, and Dean, 1979
Total			46,231	39,019	

[a]Data from these bases omitted from predictive validity analyses.

Enlisted History Tape

The Navy enlisted history tape (EHT), which is developed and maintained by the Naval Health Research Center, contains enlistment histories of all personnel who have been on active duty since 1 January 1965. EHT indices that reflect Navy job performance include total promotions, total demotions, total disciplinary infractions (e.g., leave without pay, desertion), length of service (LOS), highest pay grade achieved, rate of advancement, and AFQT score (if any). In standard Navy practice, personnel are classified on the basis of their AFQT scores into the following mental level groups:

Mental Level Group	AFQT Score	
1	93-99	Eligible for "A" schools
2	65-92	
High 3	49-64	
Low 3	31-48	Ineligible for "A" schools
High 4	21-30	
Low 4	10-20	

Identifying GENDETs

A computer algorithm was developed for use in (1) distinguishing EHT records of GENDETs from those of non-GENDETs based on characteristics normally associated with the two groups and (2) extracting the records of GENDETs for use in the study. The algorithm included the following decision steps, which were derived from conversations with Navy personnel classifiers at NAVPERSRANDCEN and the Recruit Training Command, San Diego:

1. The sample was restricted to personnel who enlisted in the Navy as E-1s.

2. Personnel with such recruit program codes as active mariner, seafarer, seafarer/airman, etc. were classified as GENDETs; those with such codes as advanced electronics, programmed school input, school guarantee program, etc. were classified as non-GENDETs.

3. Personnel who were not classifiable in Step 2 and who had AFQT scores in Mental Groups Low 3, High 4, or Low 4 (AFQT scores of 10 to 48) were classified as GENDETs.

4. Personnel who were not classifiable in Step 3 and who were nondesignated strikers as either E-2s or E-3s were classified as GENDETs.

To evaluate the algorithm, 50 personnel, whose records had been selected from the 10 data bases referred to previously, were classified as GENDETs or non-GENDETs using the algorithm and the judgment of two subject matter experts (SMEs) (a chief personnelman and a personnelman first class), working independently. A comparison of results showed that 79 percent of the personnel classified as GENDETs by both SMEs were identified as GENDETs by the algorithm. No false positives (i.e., non-GENDETs who were identified incorrectly as being GENDETs) were selected. The two SMEs reached complete agreement in classification decisions 78 percent of the time; partial agreement, 16 percent of the time; and disagreement, 6 percent of the time. These findings suggest that the data used for the study contained few, if any, non-GENDETs.

Personnel records in the 10 data bases were formed into a combined data base tape, and social security numbers (SSNs) were extracted and matched against those in the October 1979 EHT. For subjects who appeared in both the combined data base tape and the EHT, complete Navy history records were extracted and combined with predictor data taken from the original data bases. The GENDET classification algorithm was applied to the 97,677 personnel in the combined file, and records of 46,231 personnel classified as GENDETs were extracted. Of this total, the records for 39,019 GENDETs included AFQT scores (see Table 1).

The total sample of 46,231 GENDETs was used to compute predictor and criterion descriptive statistics and a criterion intercorrelation matrix. The predictive validity studies, however, were restricted to the (1) 39,019 GENDETs with AFQT scores, and (2) the eight data bases with sample sizes large enough to produce stable validity coefficients. (Data bases from the assessment center study (Siegel & Wiesen, 1977) and computerized perceptual study tests (Cory, 1976) were omitted.) A double cross-validation paradigm was used.

Predictor Variables

Predictor variables derived from operational and experimental measures are listed in Tables 2 and 3 respectively and discussed below.

Variables From Operational Measures

The operational tests used for the first seven data bases (Cory et al., 1980; Dann & Abrahams, 1973; Duffy et al., 1975) were BTB and SCT subtests, which were used by the Navy for classification purposes until 1 January 1976. The tests used for the eighth data base (Cory et al., 1979) were subtests of ASVAB Forms 5, 6, and 7, used after 1 January 1976. Although 21 operational predictor variables are listed in Table 2, the maximum number that could be used for any data base was 13 because the operational classification battery would consist of either BTB/SCT or ASVAB tests.[2]

The success chances of recruits entering the Navy (SCREEN) is a multidimensional variable that has been found (Lockman & Lurie, 1980) to predict retention and attrition of enlisted personnel. Since SCREEN was adopted operationally in October 1976, its computational formula has been changed several times. The formula used to compute SCREEN between March 1977 and October 1980 was used for this research. It included an intellectual component based on AFQT level, as well as personality and behavioral components based on age (AGE), education (ED), and marital status. Although the Navy did not use SCREEN to select any of the personnel in this research, the SCREEN scores present on the EHT were used as predictor variables for the analyses.

Variables from Experimental Measures

Experimental selection measures available for the eight data bases used in the predictive validity analyses included 20 experimental tests or questionnaires and their 125 associated scores (see Table 3). Since each set of experimental predictors was collected for only one data base, its use was limited to that data base. These predictor variables are described in greater detail in the reports describing the original research (Cory et al., 1979; Cory et al., 1980; Dann & Abrahams, 1973; Duffy et al., 1975).

---

[2]Although the ASVAB battery contains 12 subtests, only the 9 used for Navy classification of enlisted personnel were included as predictors in this study.

Table 2

Predictor Variables Derived From Operational Measures

Predictor Variable	Abbreviation	Description
Enlisted History Tape		
Armed Forces Qualification Test Score	AFQT	A measure of general mental ability, scored in centiles, based on vocabulary, arithmetic reasoning, spatial reasoning, and mechanical knowledge.
Success chances of recruits entering the Navy score	SCREEN	A score, scaled in centiles, derived from age, education, AFQT, and marital status characteristics of recruits.
Years of age at entry into recruit training	AGE	--
Years of education	ED	--
Basic Test Battery (BTB) and Special Classification Test (SCT) Form 7[a,b]		
General Classification Test	GCT	Measures word meanings and ability to reason verbally.
Arithmetic Reasoning	ARI	Measures quantitative aptitude, including mathematical reasoning and problem solving.
Clerical	CLER	Measures perceptual speed and accuracy.
Sonar Pitch Memory	SONR	Measures ability to perceive and remember small differences in tonal pitches.
Radio Code Aptitude	RADO	Measures ability to learn, remember, and use sound patterns as symbols.
Electronic Technician Selection	ETST	Measures knowledge of mathematics, science, electricity, and electronics.
Shop Practices	SHOP	Measures knowledge of tools and shop equipment.
Armed Services Vocational Aptitude Battery (ASVAB) Form 6 and 7[c]		
Numerical Operations	NO	Measures ability to perform elementary addition, subtraction, multiplication, and division.
Attention to Detail	AD	Measures ability to count quickly and accurately the number of Cs embedded in lines of Os.
Word Knowledge	WK	Measures vocabulary.
Arithmetic Reasoning	AR	Measures ability to perform arithmetic reasoning with word problems.
Mathematics Knowledge	MK	Measures knowledge of algebra, geometry, fractions, decimals, and exponents.
Electronics Information	EI	Measures knowledge of electrical and electronic components, principles, and symbols.
Mechanical Comprehension	MC	Measures basic knowledge of mechanical principles.
General Science	GS	Measures knowledge of the physical and biological sciences.
Shop Information	SI	Measures knowledge of shop practices and tool use.

[a]Scaled as Navy Standard Scores, having a mean of about 50 and a standard deviation of 10 for an unrestricted recruit population.

[b]BTB and SCT subtest scores were the operational classification tests used for seven of the eight data bases (all but BCS/Cleff) (Cory et al., 1978).

[c]ASVAB subtest scores were used as the operational classification tests for the BCS/Cleff data base.

Table 3

Predictor Variables Derived from Data Base Experimental Measures

Data Base/Experimental Measure	Experimental Predictor
Project 100,000, Phase 1 (Cory et al., 1980)	
1. Memory for Numbers	Total correct recalls Weighted total correct recalls
2. Hand Skills Test	Part 3 score minus base rate Total score
Project 100,000, Phase 2 (Cory et al., 1980)	
1. Listening Skills Test 1	Test score
2. Dominoes Test	Test score
3. Maze Test	Test score
4. Biographical Information Form	Binary rated/nonrated (SN) score Binary rated/nonrated (FN) score Binary rated/nonrated (AN) score LOS (SN) score LOS (FN) score LOS (AN) score Attrition (SN) score Attrition (FN) score Attrition (AN) score Disciplinary infractions (SN) score Disciplinary infractions (FN) score Disciplinary infractions (AN) score Supervisor's marks (SN) score Supervisor's marks (FN) score Supervisor's marks (AN) score
Project 100,000, Phase 3 (Cory et al., 1980)	
1. Listening Skills Test 2	Test score
2. Strong Vocational Interest Blank (SVIB)	Achievement scale Binary rated/nonrated (SN) score Binary rated/nonrated (FN) score Binary rated/nonrated (AN) score LOS (SN) score LOS (FN) score LOS (AN) score Attrition (SN) score Attrition (FN) score Attrition (AN) score Disciplinary infractions (SN) score Disciplinary infractions (FN) score Disciplinary infractions (AN) score Supervisor's marks (SN) score Supervisor's marks (FN) score Supervisor's marks (AN) score

Table 3 (Continued)

Data Base/Experimental Measure	Experimental Predictor
Project 100,000, Phase 4 (Cory et al., 1980)	
1. Card Pattern Test	Test score
2. Coding Test	Test score
3. Mechanical Principles Test	Test score
4. Word Finding Test	Test Score
5. Biographical Information Questionnaire (BIF) 1	Binary rated/nonrated (SN) score
	Binary rated/nonrated (FN) score
	Binary rated/nonrated (AN) score
	LOS (SN) score
	LOS (FN) score
	LOS (AN) score
	Attrition (SN) score
	Attrition (FN) score
	Attrition (AN) score
	Disciplinary infractions (SN) score
	Disciplinary infractions (FN) score
	Disciplinary infractions (AN) score
	Supervisor's marks (SN) score
	Supervisor's marks (FN) score
	Supervisor's marks (AN) score
Project 100,000, Phase 5	
1. Socio-economic Status Questionnaire	Socio-economic status scale 1
	Socio-economic status scale 2
	Socio-economic status scale 3
2. Visual Memory Test	Part 1 score
	Part 2 score
	Part 3 score
3. Following Instructions Test	Test score
	Direct recall (item 21) score
4. Biographical Information Questionnaire (BIF) 2	Scholastic motivation score
	Control score
	Self-concept score
	Binary rated/nonrated (SN) score
	Binary rated/nonrated (FN) score
	Binary rated/nonrated (AN) score
	LOS (SN) score
	LOS (FN) score
	LOS (AN) score
	Attrition (SN) score
	Attrition (FN) score
	Attrition (AN) score
	Disciplinary infractions (SN) score
	Disciplinary infractions (FN) score
	Disciplinary infractions (AN) score
	Supervisor's marks (SN) score
	Supervisor's marks (FN) score
	Supervisor's marks (AN) score

Table 3 (Continued)

Data Base/Experimental Measure	Experimental Predictor
Navy Vocational Interest Inventory (NVII) (Dann & Abrahams, 1973)	Interest in QM rating score
	Interest in ST rating score
	Interest in ET rating score
	Interest in RM rating score
	Interest in SK rating score
	Interest in CS rating score
	Interest in EN rating score
	Interest in BT rating score
	Interest in EM rating score
	Interest in EO rating score
	Interest in AO rating score
	Interest in HM rating score
	Binary rated/nonrated (SN) score
	Binary rated/nonrated (FN) score
	Binary rated/nonrated (AN) score
	LOS (SN) score
	LOS (FN) score
	LOS (AN) score
	Attrition (SN) score
	Attrition (FN) score
	Attrition (AN) score
	Disciplinary infractions (SN) score
	Disciplinary infractions (FN) score
	Disciplinary infractions (AN) score
	Supervisor's marks (SN) score
	Supervisor's marks (FN) score
	Supervisor's marks (AN) score
Gates-MacGinitie Reading Test (Form D) (Duffy et al., 1975)	Vocabulary grade level score
	Reading comprehension grade level score
	Total grade level score
BCS/Cleff Inventory (Cory et al., 1979)	Overall adjustment score
	Success in boot camp score
	Success in BT "A" School score
	Success in ET "A" School score
	Success in HM "A" School score
	Success in MM "A" School score
	Success in OS "A" School score
	Success in BT job score
	Success in ET job score
	Success in HM job score
	Success in MM job score

As shown in Table 3, five of the eight data bases included personal questionnaires, either biographical data forms or interest inventories (the SVIB or the NVII). From each of these data bases, scales were developed for the three apprenticeship groups, SN, FN, and AN, using as test construction samples personnel whose SSNs ended in 1, 3, 5, 0, or 8. A key construction technique was used to select the most predictive questions for each of five criteria: rated/nonrated (BIRTD), behavioral record (BEHR), overall performance (OVER), years of service (YRSERV), and attrition (ATTR). Together with scores for the other predictors, these scales were cross-validated on holdout samples.

Criteria

Initially, the following 10 job performance criteria were selected for evaluation:

1. Not recommended for reenlistment (NTREC)--A binary variable coded 1 for any person who has ever not been recommended for reenlistment in the Navy.

2. Total demotions (TDEMO)--An enlisted person's number of demotions in the Navy.

3. Behavioral record (BEHR)--The weighted sum of disciplinary actions, derived as follows: (2 x total demotions) + (2 x total desertions) + total unauthorized absences (UA/AWOLS).

4. Attrition (ATTR)--A binary variable indicating receipt of an attrition discharge (for disability, dependency, unsuitability, or unfitness) in contrast to normal completion of an enlistment (honorable discharge, expiration of enlistment, released inactive duty, fleet reserve).

5. Overall performance (OVER)--Supervisors' rating of overall job performance. The supervisory evaluation question requested that the supervisor compare each person's overall performance to that of other personnel in the same rate. Performances in the bottom 20 percent, lower 20 to 40 percent, middle 20 percent, upper 20 to 40 percent, and top 20 percent were coded 1, 2, 3, 4, and 5 respectively.

6. Days to achieve E-4 (DAE4)--The number of days a person required to achieve rated status.

7. Highest pay grade (HIPG)--The highest pay grade a person achieved during Navy service.

8. Rated/nonrated (BIRTD)--A binary variable coded to indicate achievement of rated status (E-4 and above) during a Navy career.

9. Years of service (YRSERV)--Total years of service from boot camp to discharge.

10. Total promotions (TOTPRO)--The number of promotions received in a Navy career.

Nine of the criterion variables (all but OVER) were derived from the EHT. These criteria range in content from BEHR and TDEMO, which were heavily influenced by behavioral components, to such variables as BIRTD, TOTPRO, and HIPG, which substantially reflect job performance. Data for OVER were available on only six of the eight data bases.

The scores for these 10 criteria were computed and intercorrelated, along with scores for AFQT. Results are presented in Table 4. As shown, the coefficients for HIPG, BIRTD, YRSERV, and TOTPRO were highly intercorrelated, ranging from .52 to .89, with five of the six being greater than .70. DAE4 was relatively independent of three of these four criteria (all but BIRTD), its correlations with them ranging from .11 to .27. BEHR correlated .79 with TDEMO, one of its constituents, but otherwise the other four enlisted history criteria (NTREC, TDEMO, BEHR, ATTR) had relatively low intercorrelations, ranging from .04 to .27. OVER had low correlations with all nine career outcome variables, with the largest being .17. In fact, OVER was not as accurate a predictor of career outcome variables as was AFQT. Its low relationship with other criteria indicates that supervisors' marks were of little use as job performance criteria for the GENDETs included in this research. Mental ability as measured by AFQT scores had moderately high association with BIRTD, HIPG, and DAE4 and low relationships with ATTR and BEHR.

Because of the high correlations noted above, the following five criteria, representing the major sources of variation, were selected for the predictive validity analyses to simplify and clarify the presentation:

1. Behavioral record (BEHR), a summary measure of disciplinary infractions.

2. Highest pay grade (HIPG) and rated/nonrated (BIRTD), both of which measure career performance.

3. Days to achieve E-4 (DAE4), a measure of rate of advancement.

4. Overall performance (OVER), supervisors' evaluation of overall performance.

Analysis

The percentages of rated GENDETs in each AFQT mental level were computed by rating. Intercorrelations among the criteria were computed and analyzed.

Multiple Regression

Subjects from the eight large data bases were assigned randomly into two weight development samples, based on the last digits of their SSNs. Sample 1 included persons whose SSNs ended in 1, 3, 5, 0, and 8; and Sample 2, in 2, 4, 6, 7, and 9. The statistical programs for social sciences (SPSS) REGRESSION program was then used to select the most predictive set of variables, using a hierarchical stepwise multiple regression program employing the accretion paradigm. The predictor composites were cross-validated on holdout samples using both multiple regression and unit weights. (The latter were computed by adding the z-scores of the variables selected for the multiple regression run in the weight development step.) Because of the similarity of the cross-validated coefficients for multiple regression and unit-weighted composites and because multiple-regression weights are impractical to use operationally, the validity analyses reported here used the unit-weighted composites. For these computations, predictor selection was restricted initially to operational variables. After a maximally predictive set of operational variables had been selected, selection was made from experimental variables. Two sets of variables were used for regression analyses:

1. The tests in the BTB or the ASVAB (as appropriate), plus AFQT, SCREEN, ED, and AGE.

Table 4

Intercorrelations Among 10 Job Performance Criteria and AFQT

	NTREC	TDEMO	BEHR	ATTR	OVER	DAE4	HIPG	BIRTD	YRSERV	TOTPRO	AFQT	Mean	SD	N
NTREC[a]	--	.04**	.13**	.17**	-.08**	.08**	-.37**	-.32**	-.26**	-.31**	-.26**	.27	.44	46,231
TDEMO[a]			.79**	.15**	.05**	.07**	-.00	.02**	.03**	.18**	.01*	.22	.50	46,231
BEHR[a]				.27**	-.07**	.10**	-.14**	-.08**	-.03**	.01**	-.06**	.75	1.60	46,231
ATTR[a]					-.08**	-.13**	-.42**	-.30**	-.42**	-.38**	-.07**	.31	.46	40,806
OVER[b]						-.02	.17**	.12**	.09**	.13**	.06**	3.48	1.18	5,625
DAE4[a]							-.17**	--	.27**	.11**	-.41**	872.76	439.88	20,159
HIPG								.71**	.77**	.89**	.36**	3.62	1.21	46,231
BIRTD									.52**	.74**	.25**	.44	.50	46,231
YRSERV										.74**	.17**	3.69	2.51	46,231
TOTPRO											.24**	2.49	1.19	46,231
AFQT												56.45	20.85	39,019

Note. All correlations are Pearson product-moment coefficients.

[a]Variable was negatively scaled: high scores are undesirable; and low scores, desirable.

[b]Data were collected for only a portion of the sample.

*p<.05.
**p<.001.

2. All variables in Set 1 plus the experimental variables associated with each data base (see Table 3).

The following rules were used when selecting variables for predictor composites:

1. No more than five variables were selected for a composite.

2. Aptitude variables that were weighted negatively by the program were not included because negatively weighted aptitude variables could never be used operationally for Navy personnel selection or classification.

Differential Predictability of Apprenticeship Subgroups

To determine whether developing separate prediction equations for the SN, FN, and AN apprenticeship ratings would be a significant improvement over the use of a single prediction equation for the whole group, multiple regression analyses (Bottenberg & Ward, 1963) were performed separately within each of the eight large data bases. Two sets of multiple regression variables were selected for each criterion in each data base: (1) the best set of predictors for the whole data base, and (2) the most predictive set for each of the three subgroups (a total of three sets). To determine whether differential prediction (separate predictor composites for SN, FN, and AN) significantly improved predictive accuracy over a single predictor composite for the whole group, squared errors of prediction on cross-validation samples were computed for each method and tested statistically. Failure of a weighted composite coefficient for the three subgroups to be significantly higher than that for the whole group would indicate that the subgroup equations were no more accurate than the equation for the total group. Therefore, the best policy would be to compute one prediction equation for the total group instead of separate equations for the three subgroups.

RESULTS

GENDET Performance and Advancement

Predictive validity analysis using five criteria that represented major sources of criterion variation[3] produced consistent relationships between mental ability as measured by AFQT and BIRTD, DAE4, and YRSERV. Table 5 shows that, as AFQT mental ability levels lowered, the percentage of personnel who became rated (BIRTD) decreased, the time required to become rated (DAE4) increased, and the LOS (YRSERV) decreased. These effects were very large. For instance, the percentage of Mental Group 1 personnel who became rated (BIRTD) was nearly four times that of Mental Group High 4, nearly twice that of Low 3, and 50 percent greater than High 3. Similar large differences were evident in time required to achieve rated status (DAE4) and LOS (YRSERV). In contrast, BEHR and ATTR had little or no relationship to mental ability.

Percentage of Enlistment Period Served at E-4 and Above

On the average, the percentage of a 4-year enlistment during which GENDETs served at the E-3 pay grade or higher varied little from one AFQT mental level group to another (see Table 6). In contrast, a substantial monotonic increasing relationship existed

[3] These are not the five criteria listed earlier. For purposes of this analysis, YRSERV and ATTR were substituted for HIPG and OVER because they were more descriptive of performance by mental level group.

Table 5

Mean Job Performance of GENDETs by AFQT Mental Level Group

Criterion	AFQT Mental Level Group						
	1	2	High 3	Low 3	High 4	Low 4	N[a]
BIRTD (percentage)	77	62	51	42	20	73	39,019
DAE4 (mean number of days)	580	726	929	1,026	1,301	1,487	19,487
YRSERV (mean number of years)	5.13	4.44	3.96	3.66	2.97	7.59	39,019
ATTR (percentage)	23	26	32	41	18	34	33,868
BEHR	.55	.66	.85	.99	.58	.83	39,019

[a]Sample sizes are smaller than in Table 4 because of missing AFQT scores.

Table 6

Average Time Spent at or Above Three Pay Grades During 4-Year Enlistment

AFQT Mental Level Group	N	Average Percentage Days Worked at Pay Grade		
		E-3	E-4	E-5
1	1,629	72	46	9
2	11,893	67	31	3
High 3	10,332	60	18	-
Low 3	11,350	51	12	-
High 4	3,522	48	2	-
Low 4	293	67	-	-
Total	39,019	59	20	1

between mental group level and percentage of time spent as E-4 and E-5. Personnel with higher mental abilities spent a much higher percentage of their overall enlistment period at E-4 and above than did those with lower mental abilities. For instance, Mental Group 1s, on the average, spent 46 percent of their enlistment period as E-4s, compared with 12, 2, and 0 percent for Low 3s, High 4s, and Low 4s.

If time spent as a rated person is of greater value to the Navy than time spent on the job in partial performance or training, these data clearly demonstrate the importance of mental level in selecting enlisted personnel for GENDET assignments.

Assignment to Navy Occupational Specialties (Ratings)

Of the 39,019 GENDETs with AFQT scores, 51 percent (19,587) were assigned to Navy occupational specialties or ratings. Out of 58 ratings, 4--boatswain's mate (BM), machinist's mate (MM), radioman (RM), and electronics technician (ET)--each received 5 percent or more of the rated GENDETs. Nine other ratings had between 3 and 5 percent of the total. These 13 ratings (22.4% of the 58 ratings) combined received 56.4 percent of the rated GENDETs.

The percentages of GENDETs in each AFQT mental level group in various Navy ratings provide a rough indication of the mental levels required for success in each. For example, as shown in Table 7, 61.8 percent of personnel in the BM rating were in Mental Levels 1, 2, or High 3; therefore, the mental level concentration for BM was coded school-eligible (SE). In contrast, the mental level concentration for mess management specialist (MS) was coded as non-school-eligible (NSE), because 86.8 percent of its rated personnel were in Mental Levels Low 3, High 4, or Low 4.

NSE personnel constituted 28 percent of GENDETs who became rated, appearing in all 58 occupational specialties studied. Nine ratings had greater concentrations of NSE than SE personnel. Of these, three ratings--MS, hospital corpsman (HM), and dental technician (DT)--included 91 percent of NSE personnel. The other six ratings included percentages of the total sample ranging from .0 to .3. In the three larger ratings with NSE concentrations, the NSE personnel who became rated were predominantly Low 3s.

These data suggest that there are few occupational specialties for which the job requirements are appropriate for NSE personnel, and that these positions are primarily appropriate for Low 3s, although some High 4s and Low 4s can become rated.

Predictive Validity Analyses

Descriptive statistics for the variables selected for predictive composites are shown in Table 8. For the AFQT, BTB, and SCT, means were near the midpoint of the range (50) and the standard deviations (SDs) were close to those for full-range samples (20 for AFQT, 10 for BTB and SCT); the distributions were normal. In contrast, the means and SDs for the two ASVAB scores indicate that these distributions were markedly below normal and were highly restricted in range, undoubtedly because the majority of subjects in the BCS/Cleff study were NSE personnel with ASVAB scores ranging from 42 to 48.

Table 7

GENDETs of Each Mental Level Group by Navy Rating

Rating	% of Total N	% of Total %	Mental Level (ML): 1 N	1 %	2 N	2 %	High 3 N	High 3 %	Low 3 N	Low 3 %	High 4 N	High 4 %	Low 4 N	Low 4 %	ML Concentration SE	ML Concentration NSE	Apprentice Field
Boatswain's mate	1451	7.4	11	.8	292	20.1	593	40.9	480	33.1	66	4.5	9	.6	X	--	SN
Machinist's mate	1081	5.5	66	6.1	373	34.5	351	32.5	250	23.1	32	3.0	9	.8	X	--	FN
Radioman	1004	5.1	31	3.1	374	37.3	264	26.3	313	31.2	19	1.9	3	.3	X	--	SN
Electronics technician	977	5.0	264	27.0	627	64.2	68	7.0	17	1.7	1	.1	0	.0	X	--	SN
Mess management specialist	846	4.3	9	1.1	89	10.5	161	19.9	363	42.9	150	17.7	74	8.7	--	X	SN
Engineman	829	4.2	30	3.6	278	33.5	303	36.6	180	21.7	31	3.7	7	.8	X	--	FN
Yeoman	792	4.0	29	3.7	225	28.4	290	36.6	215	27.1	24	3.0	9	1.1	X	--	SN
Boiler technician	767	3.9	8	1.0	226	29.5	284	37.0	222	28.9	24	3.1	3	.4	X	--	FN
Aviation electronics technician	761	3.9	123	16.2	542	71.2	71	9.3	25	3.3	0	.0	0	.0	X	--	AN
Hull maintenance technician	698	3.5	11	1.6	207	29.6	222	31.8	219	31.4	35	5.0	4	.6	X	--	FN
Aviation structural mechanic	658	3.4	20	3.0	237	36.0	156	23.7	207	31.2	29	4.4	11	1.7	X	--	AN
Aviation ordnanceman	610	3.1	88	14.4	353	57.9	100	16.4	58	9.5	9	1.5	2	.3	X	--	AN
Electrician's mate	606	3.1	49	8.1	256	42.2	152	25.1	107	17.7	35	5.8	7	1.2	X	--	FN
Storekeeper	532	2.7	8	1.5	138	25.9	155	29.1	167	31.4	50	9.4	14	2.6	--	X	SN
Gunner's mate	502	2.6	15	3.0	164	32.7	169	33.7	148	29.5	6	1.2	0	.0	X	--	SN
Aviation machinist's mate	496	2.5	24	4.8	194	39.1	117	23.6	142	28.6	11	2.2	8	1.6	X	--	AN
Ship's serviceman	485	2.5	2	.4	81	16.7	154	31.8	187	38.6	48	9.9	13	2.7	--	X	SN
Aviation electrician's mate	471	2.4	17	3.6	318	67.5	99	21.0	32	6.8	3	.6	2	.4	X	--	AN
Quartermaster	470	2.4	9	1.9	149	31.7	149	31.7	145	30.9	16	3.4	2	.4	X	--	SN
Aviation boatswain's mate	464	2.4	7	1.5	136	29.3	175	37.7	131	28.2	8	1.7	7	1.5	X	--	AN
Operations specialist	442	2.3	27	6.1	211	47.7	119	26.9	81	18.3	4	.9	0	.0	X	--	SN
Hospital corpsman	441	2.3	8	1.8	46	10.4	103	23.4	268	60.8	15	3.4	1	.2	--	X	HN
Sonar technician	430	2.2	65	15.1	278	64.7	63	14.7	24	5.6	0	.0	0	.0	X	--	SN
Communication technician (now cryptologic technician)	430	2.2	65	15.1	229	53.3	79	18.4	56	13.0	1	.2	0	.0	X	--	SN
Interior communications electrician	407	2.1	62	15.2	207	50.9	79	19.4	55	13.5	4	1.0	0	.0	X	--	FN
Fire control technician	397	2.0	60	15.1	263	66.2	48	12.1	25	6.3	1	.3	0	.0	X	--	SN
Torpedoman's mate	257	1.3	12	4.7	114	44.4	73	28.4	58	22.6	0	.0	0	.0	X	--	SN
Personnelman	235	1.2	9	3.8	85	36.2	80	34.0	53	22.6	7	3.0	1	.4	X	--	SN
Aviation storekeeper	192	1.0	2	1.0	41	21.4	61	31.8	65	33.9	11	5.7	12	6.3	--	X	AN
Postal clerk	184	.9	1	.5	36	19.6	77	41.8	53	28.8	15	8.2	2	1.1	X	--	SN

[a]Categorized as school-eligible (SE) if ≥ 60 percent of the rated personnel were in mental levels 1, 2, or High 3, or as non-school eligible (NSE) if ≥ 60 percent were Low 3s, High 4s, or Low 4s.

Table 7 (Continued)

Rating	% of Total N	% of Total %	Mental Level (ML) 1 N	1 %	2 N	2 %	High 3 N	High 3 %	Low 3 N	Low 3 %	High 4 N	High 4 %	Low 4 N	Low 4 %	ML Concentration SE	ML Concentration NSE	Apprentice Field
Aviation maintenance administrationman	176	.9	1	.6	52	29.5	61	34.7	55	31.3	5	2.8	2	1.1	X	--	AN
Machinery repairman	155	.8	9	5.8	49	31.6	55	35.5	33	21.3	7	4.5	2	1.3	X	--	FN
Aviation support equipment technician	141	.7	21	14.9	51	36.2	43	30.5	19	13.5	5	3.5	2	1.4	X	--	AN
Disbursing clerk	126	.6	2	1.6	32	25.4	29	23.0	53	42.1	6	4.8	4	3.2	--	X	SN
Dental technician	99	.5	0	.0	12	12.1	20	20.2	63	63.6	3	3.0	1	1.0	--	X	DN
Photographer's mate	97	.5	12	12.4	46	47.4	23	23.7	14	14.4	2	2.1	0	.0	X	--	AN
Air controlman (now air traffic controller)	93	.5	12	12.9	49	52.7	20	21.5	11	11.8	1	1.1	0	.0	X	--	AN
Aviation antisubmarine warfare operator	91	.5	10	11.0	70	76.9	5	5.5	5	5.5	1	1.1	0	.0	X	--	AN
Tradevman (rating being phased out)	88	.4	16	18.2	60	68.2	9	10.2	3	3.4	0	.0	0	.0	X	--	AN
Data processing technician	83	.4	7	8.4	34	41.0	31	37.3	11	13.3	0	.0	0	.0	X	--	SN
Aircrew survival equipmentman	71	.4	7	9.9	38	53.5	19	26.8	7	9.9	0	.0	0	.0	X	--	AN
Lithographer	65	.3	1	1.5	11	16.9	28	43.1	23	35.4	2	3.1	0	.0	X	--	SN
Aerographer's mate	59	.3	5	8.5	34	57.6	10	16.9	10	16.9	0	.0	0	.0	X	--	AN
Equipment operator	53	.3	2	3.8	7	13.2	11	20.8	28	52.8	3	5.7	2	3.8	--	X	CN
Builder	42	.2	0	.0	4	9.5	8	19.0	28	66.7	1	2.4	1	2.4	--	X	FN
Musician	32	.2	2	6.3	15	46.9	9	28.1	5	15.6	0	.0	1	3.1	X	--	SN
Utilitiesman	26	.1	0	.0	0	.0	3	11.5	22	84.6	1	3.8	0	.0	--	X	CN
Instrumentman	24	.0	4	16.7	12	50.0	5	20.8	3	12.5	0	.0	0	.0	X	--	SN
Illustrator draftsman	21	.1	1	4.8	7	33.3	9	42.9	3	14.3	1	4.8	0	.0	X	--	SN
Construction electrician	19	.1	0	.0	6	31.6	2	10.5	11	57.9	0	.0	0	.0	--	X	CN
Construction mechanic	18	.1	1	5.6	3	16.7	2	11.1	10	55.6	2	11.1	0	.0	--	X	CN
Mineman	16	.1	0	.0	8	50.0	1	6.3	7	43.8	0	.0	0	.0	X	--	SN
Steelworker	14	.1	0	.0	2	14.3	0	.0	12	85.7	0	.0	0	.0	--	X	CN
Opticalman	13	.1	1	7.7	8	61.5	3	23.1	1	7.7	0	.0	0	.0	X	--	SN
Journalist	13	.1	1	7.7	6	46.2	3	23.1	3	23.1	0	.0	0	.0	X	--	SN
Boilermaker (now boiler technician)	13	.1	1	7.7	4	30.8	3	23.1	5	38.5	0	.0	0	.0	X	--	FN
Molder	12	.1	0	.0	5	41.7	4	33.3	2	16.7	1	8.3	0	.0	X	--	FN
Photographic intelligenceman (now intelligence specialist)	7	.0	1	14.3	3	42.9	2	28.6	1	14.3	0	.0	0	.0	X	--	AN
Engineering aide	5	.0	0	.0	0	.0	0	.0	3	60.0	2	40.0	0	.0	--	X	CN
Total	19587	100.0	1249	--	7397	--	5233	--	4795	--	698	--	215	--	45	9	

[a]Categorized as school-eligible (SE) if ≥ 60 percent of the rated personnel were in mental levels 1, 2, or High 3, or as non-school eligible (NSE) if ≥ 60 percent were Low 3s, High 4s, or Low 4s.

Table 8

Descriptive Statistics for Predictors Included in Composites

Predictor Variable	Mean	SD	Skewness	Kurtosis	N
Operational					
Enlisted History Tape:					
AFQT	56.48	20.85	.20	-1.02	39,019
Years of age at enlistment (AGE)	19.29	1.28	3.37	68.71	46,230
Years of education (ED)	11.67	1.08	-.98	4.66	26,195
SCREEN	82.46	8.10	-1.07	.62	39,588
BTB Tests:					
General Classification Test (GCT)	49.26	9.73	.00	-.26	16,924
Arithmetic Reasoning (ARI)	50.61	9.00	.06	-.39	34,337
Mechanical (MECH)	49.97	8.35	.08	-.42	34,337
Clerical (CLER)	50.33	8.49	.04	-.17	34,326
SCT:					
Sonar Pitch Memory (SONR)	52.11	9.89	.11	-.65	30,729
Electronics Technician Selection (ETST)	52.99	9.69	-.05	-.14	32,913
Shop Practices (SHOP)	52.16	8.93	-.16	-.28	31,395
Radio Code Aptitude (RADO)	54.12	9.89	.11	-.65	30,729
ASVAB Tests:					
Word Knowledge (WK)	46.16	5.43	.11	.40	1,249
Mathematics Knowledge (MK)	45.66	7.50	-.09	-.44	1,249
Experimental					
Project 100,000, Phase 3:					
SVII achievement scale (ACH)	32.67	12.98	.38	-.40	1,244
SCII scale for predicting binary rated/nonrated score (FBR1)	101.32	5.04	-.26	-.57	5,991
Project 100,000, Phase 5:					
Scale of BIF2 predicting binary rated/nonrated score (FBR2)	108.11	5.68	-.10	-.48	3,684
NVII:					
Scale for predicting LOS for FN (FLS1)	115.80	7.47	-.05	-.80	12,933
Scale for predicting LOS for SN (SLS1)	119.95	7.43	-.05	-.69	12,933
BSC/Cleff Inventory:					
Scale for predicting success in MM fleet performance (MMF)	10.61	22.85	.02	-.02	853

As might be expected from scores on tests administered during Navy enlistment, the most marked differences from a normal distribution occurred for AGE and ED: The distributional concentration for AGE was between 18 and 19 years; most personnel had 12th grade educations and were high school graduates.

Validity Coefficients

For each of the eight large data bases, predictor composites were developed separately for weight development samples 1 and 2, using the regression techniques described previously. Sample 1 predictor equations were cross-validated on those from Sample 2; and Sample 2 equations were cross-validated on those from Sample 1. The coefficients resulting from these operations are presented in Table 9. In addition, back-validity coefficients were computed by applying the unit-weighted composites from each sample to itself. Coefficients resulting from back validation are presented in Table 10.

As shown in Table 9, when the composites were cross-validated against the job performance criteria, only 9 of the 76 correlation coefficients indicated that experimental predictors had improved the validity of the composites formed from operational measures. Since only four of the nine operational-experimental coefficients provided more than .01 increase, experimental variables did not substantially increase the predictiveness of the operational variables.

Although the magnitude of the cross-validation coefficients for each criterion varied somewhat from one data base to another, the rank order of the coefficients remained fairly constant. In general, the coefficients for the BCS/Cleff sample were lower than those for the other seven data bases, perhaps because the BCS/Cleff data base was highly restricted in range of variation and was concentrated in the Low 3 mental ability range (see Table 8).

Differential Predictability for SN, FN, and AN Subgroups

Although a comprehensive series of analyses was performed to develop separate predictive composites for the SN, FN, and AN apprenticeship fields, these composites were no more accurate on cross validation than was a single composite computed for the combined group. As shown in Table 11, when weighted sums of cross-validated coefficients for SN, FN, and AN subgroups were formed and their predictiveness compared with the undifferentiated group of SNs, FNs, and ANs, the coefficients for the undifferentiated group were higher than the combined coefficients for the subgroups in 77 percent of the comparisons. The difference in scores obtained by subtracting the coefficients in Table 9 from those in Table 11 are shown in Table 12.

When the average coefficients were computed across the eight data bases, the undifferentiated group coefficient for each of the five criteria was higher than the combined coefficients for the rating subgroups. Thus, it was not possible by means of the predictors used in this research to develop a differential battery that would improve the assignment of GENDETs. Because separate predictor composites for SN, FN, and AN would not improve predictive accuracy, the remaining validity investigations were carried out using composites computed for the combined sample.

Table 9

Cross Validities of Predictor Composites for Five Job Performance Criteria in Weight Development Samples by Data Base

Criterion	Project 100,000 P1	P2	P3	P4	P5	NVII	GM[a]	BCS/Cleff[a]
Weight Development Sample 1 (N = 18,033)								
HIPG	.50***	.33***	.30*** .34***	.59***	.61***	.42***	.37***	.23***
BIRTD	.43***	.27***	.33*** .34***	.36***	.58*** .58***	.34***	.31***	.12***
BEHR	-.06**	-.19***	.07***	-.15***	-.11***	-.14***	-.28***	-.17***
DAE4	-.31***	-.25***	-.30***	-.51***	-.50***	-.25*** -.24***	-.25***	-.38*
OVER	.11*	-.06	.26***	.02	.19***	.18*** .19***	--	--
Weight Development Sample 2 (N = 18,660)								
HIPG	.52***	.38***	.47*** .48***	.55***	.49***	.37***	.43***	.24***
BIRTD	.49***	.27***	.31*** .41***	.38***	.43***	.32***	.33***	.18***
BEHR	-.07**	-.09***	-.08***	-.12***	-.10***	-.13***	-.19***	-.16***
DAE4	-.20***	-.32***	-.31***	-.54***	-.52***	-.25***	-.11*	-.07 -.27
OVER	.00	.11	.31***	.01	.16**	.22*** .26***	--	--

Note. Coefficients were produced by cross-validating composites formed from operational predictors against the five job performance criteria. In the cases where experimental predictors added to the predictiveness of a weight development composite, a second coefficient--for the combined operational-experimental predictor composite--is shown.

[a]No criterion data were available for OVER in this data base.

*p<.05.
**p<.01.
***p<.001.

Table 10

Back Validities of Predictor Composites for Five Job Performance Criteria in Weight Development Samples by Data Base

Criterion	Project 100,000					NVII	GM[a]	BCS/Cleff[a]
	P1	P2	P3	P4	P5			
				Sample 1 (N = 18,279)				
HIPG	.52***	.38***	.36*** .38***	.56***	.58***	.40***	.42***	.28***
BIRTD	.48***	.29***	.33*** .49***	.38***	.48***	.32***	.35***	.21***
BEHR	-.07***	-.09***	-.08***	-.14***	-.10***	-.13***	-.27***	-.16***
DAE4	-.26***	-.33***	-.35***	-.49***	-.47***	-.27*** -.28***	-.20***	-.37*
OVER	.05	.15	.35***	.10*	.26***	.28*** .32***	--	--
				Sample 1 (N = 18,416)				
HIPG	.50***	.34***	.48*** .49***	.58***	.55***	.41***	.41***	.25***
BIRTD	.45***	.30***	.37*** .49***	.36***	.53***	.34***	.34***	.11***
BEHR	-.06***	-.19***	-.08***	-.14***	-.13***	-.14***	-.19***	-.17***
DAE4	-.33***	-.27***	-.38***	-.54***	-.55***	-.25***	-.24***	-.46** -.52**
OVER	-.13*	-.03	.24***	.21**	.16**	.18*** .21***	--	--

Note. Coefficients were produced by cross-validating composites formed from operational predictors against the five job performance criteria. In the cases where experimental predictors added to the predictiveness of a weight development composite, a second coefficient is shown.

[a]No criterion data were available for OVER in this data base.

*p<.05.
**p<.01.
***p<.001.

Table 11

Cross Validities of Predictor Composites for SN, FN, and AN Subgroups for Weight Development Sample 1 by Data Base (N = 17,053)

	Project 100,000							
Criterion	P1	P2	P3	P4	P5	NVII	GM	BCS/Cleff
				SN				
HIPG	.53***	.33***	.39***	.62***	.56***	.35***	.42***	.27***
BIRTD	.42***	.26***	.40***	.37***	.55***	.29***	.37***	.03
BEHR[a]	-.09**	--	.05**	.00	.07	-.05***	-.20***	-.04
DAE4	-.35***	-.22***	-.38***	-.53***	-.60***	-.35***	-.25***	-.34
OVER[a]	.03	--	.22***	--	.28**	.08	--	--
				FN				
HIPG	.33***	.32***	.43***	.49***	.58***	.33***	.24***	.27***
BIRTD[a]	.42***	.23***	.26***	.27***	.54***	.18***	.20***	--
BEHR	-.07	-.25***	.07**	-.15***	.04	-.17***	-.15**	-.29***
DAE4[a]	--	-.14**	-.18**	-.44**	-.38***	-.31***	-.16***	--
OVER[a]	-.05	--	-.02	.17	.23	.08	--	--
				AN				
HIPG	.35***	.30***	.45***	.62***	.57***	.48***	.34***	.21**
BIRTD[a]	.44***	.22***	.24***	.33***	.57***	.28***	.20***	--
BEHR[a]	-.13***	--	.06	-.18***	.17***	-.07**	-.07	-.11
DAE4[a]	--	-.22***	-.47***	-.51***	-.30***	-.07**	-.40***	.10
OVER[a]	--	--	.16**	--	.27*	.06	--	--
				Combined SN, FN, and AN				
HIPG	.43	.32	.42	.60	.57	.38	.36	.27
BIRTD[a]	.42	.26	.32	.34	.55	.27	.31	--
BEHR[a]	-.09	--	.06	-.05	.11	-.09	-.17	-.10
DAE4[a]	--	-.19	-.36	-.52	-.44	-.26	-.28	--
OVER[a]	--	--	.17	--	.26	.07	--	--

Note. The total N for this table is smaller than for other tables that present analysis of Sample 1 because the most predictive variables for many rating subgroups had smaller Ns than the variables for the combined groups.

[a]Some data bases did not have statistically significant weight development composites.

*p<.05.
**p<.01.
***p<.001.

Table 12

Difference Scores Obtained by Subtracting Coefficients in Table 9 From Those in Table 11

	Project 100,000									
Criterion	P1	P2	P3	P4	P5	NVII	GM	BCS/Cleff	Sum	Avg
HIPG[a]	-.07	-.01	.08	.01	-.04	-.04	-.01	.04	-.04	-.005
BIRTD[a,b]	-.01	-.01	-.02	-.02	-.03	-.07	.00	--	-.16	-.023
BEHR[b,c]	-.03	--	.13	.11	.22	.05	.11	.07	.66	.094
DAE4[b,c]	--	.06	-.06	-.01	.06	-.01	-.03	--	.01	.002
OVER[a,b]	--	--	-.09	--	.07	-.12	--	--	-.14	-.046
	Percent Discriminating in a Negative Direction									
	67	100	60	50	80	80	75	50	----	72

[a]A negative entry indicates that the coefficient for the undifferentiated total group was larger than the average coefficient for the combined subgroups.

[b]The weight development regression runs for some data bases did not produce composites for this criterion that were statistically significant.

[c]Since this variable was scaled negatively, a negative entry indicates that the coefficient for the undifferentiated total group was smaller than the average coefficient for the combined subgroups.

Stability of Coefficients

To provide more stable estimates of the predictability of the five job performance criteria, weighted means and SDs of the cross-validation and back-validation coefficients (Tables 9 and 10) were computed using formulas recommended by Schmidt and Hunter (1977). Coefficients for the BCS/Cleff data base were not included because they were not representative for a full range sample.

As shown in Table 13, the predictability of job performance criteria derived from Navy records (HIPG and BIRTD) was generally much higher and somewhat more stable than predictability of supervisors' ratings of overall performance (OVER). When the coefficients for BIRTD were adjusted to compensate for the restriction in magnitude occurring because it was a binary variable, the average coefficients for it and HIPG were equal (biserial r = .43), while the average coefficient for OVER was .15.

Table 13

Weighted Means and Standard Deviations of Predictive Validity of Cross- and Back-validation Coefficients

Criterion	Weighted Sample 1 (N = 18,033)				Weighted Sample 2 (N = 18,660)				Average Cross-val.	
	Cross-val.		Back-val.		Cross-val.		Back-val.			
	M	SD	M	SD	M	SD	M	SD	M	SD
HIPG	.43	.07	.43	.07	.43	.09	.45	.07	.43	.08
BIRTD	.34	.05	.35	.06	.35	.07	.36	.05	.345[a]	.06
BEHR	-.12	.03	-.12	.04	-.13	.05	-.13	.03	-.125	.04
DAE4	-.28	.1	-.30	.07	-.29	.09	-.32	.01	-.285	.095
OVER	.16	.1	.22	.1	.14	.08	.17	.09	.15	.09

[a]When adjusted to compensate for the fact that the criterion is a binary variable (point-biserial coefficient converted to its biserial equivalent), the average cross-validation coefficient is .43.

The stability of the predictive validity coefficients is shown by the small average shrinkage between cross-validation and back-validation coefficients. The coefficients for three of the cross- and back-validation pairs (HIPG Sample 1, and BEHR Samples 1 and 2) were equal; the shrinkage for the other five pairs (BIRTD and DAE4 Samples 1 and 2, and HIPG Sample 2) ranged from .01 to .03. The shrinkage for OVER was somewhat higher--from .03 to .06.

Most Effective Predictors

Although 125 experimental predictors were evaluated, nearly all of the predictiveness of the composites was supplied by variables that were available operationally, either personnel classification test scores, the SCREEN score, or biographical variables such as AGE and ED. Table 14 presents the variables most highly predictive of each criterion in the eight data bases. As shown, HIPG, BIRTD, DAE4, and OVER were best predicted by the SCREEN score and scores from the BTB (particularly GCT and ARI). The best predictors of BEHR were SCREEN and AGE.[4] Experimental variables that were predictive were primarily biographical scales developed from the SVIB, the NVII, or the BCS/Cleff questionnaire.

[4]In some instances, AGE was selected as a predictor by the multiple regression program even though it is included in SCREEN, a variable also selected by the program. The effect of this double selection was to double-weight the age variable in the predictor composites used for several of the criteria in the study.

Table 14

Variables Most Predictive of Five Job Performance Criteria by Data Base

Criterion	Project 100,000 P1	P2	P3	P4	P5	NVII	GM	BCS/C	Variables most frequently selected by SPSS multiple regression program
HIPG	GCT(2) MECH(2) ARI(2)	SHOP(2) MECH(2) CLER(1) AFQT(1) SCREEN(1) AGE(1)	ARI(2) FBR2(2) MECH(1) SHOP(1) SCREEN(1) AGE(1)	GCT(2) SCREEN(2) SHOP(1) MECH(1)	GCT(2) SCREEN(1) MECH(1) SHOP(1) CLER(1) ARI(1)	SCREEN(2) MECH(2) SHOP(1) ARI(1)	SCREEN(2) ARI(2) ETST(2) AGE(2) ED(1) CLER(1)	SCREEN(2) MK(1) ED(1)	SCREEN(9) MECH(9) ARI(8) GCT(6) SHOP(6)
BIRTD	GCT(2) MECH(2) ARI(1) CLER(1)	SHOP(2) MECH(1) ARI(1) AFQT(1) SONR(1) AGE(1)	ARI(2) FBR1(1) SHOP(1) ACH(1)	GCT(2) SCREEN(2)	GCT(2) MECH(2) ARI(1) SCREEN(1) FBR1(1)	SCREEN(2) ARI(2) SHOP(1)	SCREEN(2) ARI(2) ETST(1) CLER(1) ED(1) AGE(1)	ED(2) MK(1)	ARI(8) GCT(6) MECH(5) SCREEN(5) SHOP(4)
BEHR	ARI(2)	SCREEN(2) AGE(2)	SCREEN(2) AGE(2)	SCREEN(2) AGE(2) CLER(1)	SCREEN(1) AGE(2)	SCREEN(2) AGE(2)	SCREEN(2) ETST(1) ED(1)	SCREEN(2)	SCREEN(14) AGE(10) ARI(2)
DAEA	SHOP(2) ARI(1)	SCREEN(2) AGE(2) SHOP(1) ARI(1) MECH(1)	GCT(2) ARI(1)	GCT(2) MECH(1) AFQT(1)	GCT(2) MECH(2) ARI(1) AFQT(1) CLER(1)	AFQT(2) ED(2) SLSI(1)	GCT(1) ARI(1) CLER(1) AGE(1)	MMF(2) WK(1)	GCT(6) ARI(5) MECH(4) AGE(3)
OVER	GCT(1) RADO(1) SONR(1) AGE(1)	SCREEN(1) MECH(1) ARI(1) ETST(1) RADO(1)	GCT(2) MECH(1) ED(1)	GCT(2) RADO(1)	SCREEN(2) AGE(1) ARI(1)	MECH(2) SCREEN(1) AGE(1) ED(1) FLSI(1)			GCT(4) SCREEN(4) MECH(4) AGE(3)

Note. The number of times a variable was selected by the multiple regression program is shown in parenthesis following the variable acronym.

CONCLUSIONS

A major objective in assigning personnel to GENDET positions is to provide on-the-job training to people who are not well suited for training by means of the formal academic instruction given in "A" schools. The findings of the present research indicate that SE peronnel, who have an advantage in assimilating "A" school training, also have an advantage in assimilating on-the-job training.

Greater percentages of SE than NSE GENDETs achieved rated status. SE personnel required less time to achieve rated status on the average and, consequently, spent much larger percentages of their enlistment working at the full performance level. They also had greater average LOS than did NSE personnel. However, although the service characteristics of NSE personnel were less favorable than those of SE personnel, 38 percent of NSEs achieved E-4 and above, filling 24 percent of the rated billets. Low 3s constituted 84 percent of the NS personnel who achieved E-4 and above, compared with 12 and 4 percent for High and Low 4s.

Most subjects in this research enlisted during the late 1960s and early 1970s, when larger proportions of Mental Groups 1 and 2 were enlisted in the Navy than after the advent of the all-volunteer force (AVF). Therefore, it is likely that somewhat higher proportions of High and Low 3s are now becoming rated.

The findings of this research indicate that, so long as there is an excess of applicants to GENDET billets, the quality and subsequent job performance of GENDET personnel can be improved by using predictor composites to select them.

RECOMMENDATIONS

1. Future validations against GENDET job performance should use DAE4 and either HIPG or BIRTD as criteria instead of OVER.

2. In future research using job performance criteria to select GENDETs, a single predictive composite for the combined SN, FN, and AN ratings should be computed.

3. To maximize the percentage of GENDETs who achieve rated status, selection should be based on ASVAB scores and SCREEN score. To maximize the behavioral characteristics of GENDETs, selection should be based on SCREEN score and age.

REFERENCES

Bottenberg, R. A., & Ward, J. H. Applied multiple linear regression (PRL-TDR-63-6). Lackland Air Force Base, TX, March 1963.

Cory, C. H. An evaluation of computerized tests as predictors of job performance: II. Differential validity for global and job element criteria (NPRDC Tech. Rep. 76-28). San Diego: Navy Personnel Research and Development Center, January 1976. (AD-A020 867)

Cory, C. H. The assignment of general detail personnel in the Navy: Fleet follow-up of personnel appraised in a technical classification assessment center pilot study (NPRDC Tech. Note 82-23). San Diego: Navy Personnel Research and Development Center, July 1982.

Cory, C. H., Neffson, N. E., & Dean, R. D. A controlled experimental application of the BCS/Cleff technique for the Navy (NPRDC Spec. Rep. 79-26). San Diego: Navy Personnel Research and Development Center, September 1979.

Cory, C. H., Neffson, N. E., & Rimland, B. Validity of a battery of experimental tests in predicting performance of Navy Project 100,000 personnel (NPRDC Tech. Rep. 80-35). San Diego: Navy Personnel Research and Development Center, September 1980. (AD-A091 243)

Dann, J. E., & Abrahams, N. M. Occupational scales of the Navy Vocational Interest Inventory: I. Development (NPRDC Tech. Rep. 74-4). San Diego: Navy Personnel Research and Development Center, October 1973. (AD-770 558)

Duffy, T. M., Carter, J. F., Fletcher, J. D., & Aiken, E. G. Language skills: A prospectus for the naval service (NPRDC Spec. Rep. 76-3). San Diego: Navy Personnel Research and Development Center, October 1975.

Duffy, T. M., & Nugent, W. A. Reading skills levels in the Navy (NPRDC Tech. Rep. 78-19). San Diego: Navy Personnel Research and Development Center, April 1978. (AD-A054 859)

Hiatt, C. M., & Sims, W. H. Armed Services Vocational Aptitude Battery (ASVAB) and job performance (CNA Working Paper 80-3121). Alexandria, VA: Center for Naval Analyses, November 1980.

Lockman, R. F., & Lurie, P. M. A new look at success chances of recruits entering the Navy (SCREEN) (CRC 425). Alexandria, VA: Center for Naval Analyses, February 1980.

Lurie, P. M. Relating enlistment standards to job performance: A pilot study for two Navy ratings (CNA Memorandum 81-0048). Alexandria, VA: Center for Naval Analyses, January 1981.

Schmidt, F. L., & Hunter, J. E. Development of a general solution to the problem of validity generalization. Journal of Applied Psychology, 1977, 62, 529-540.

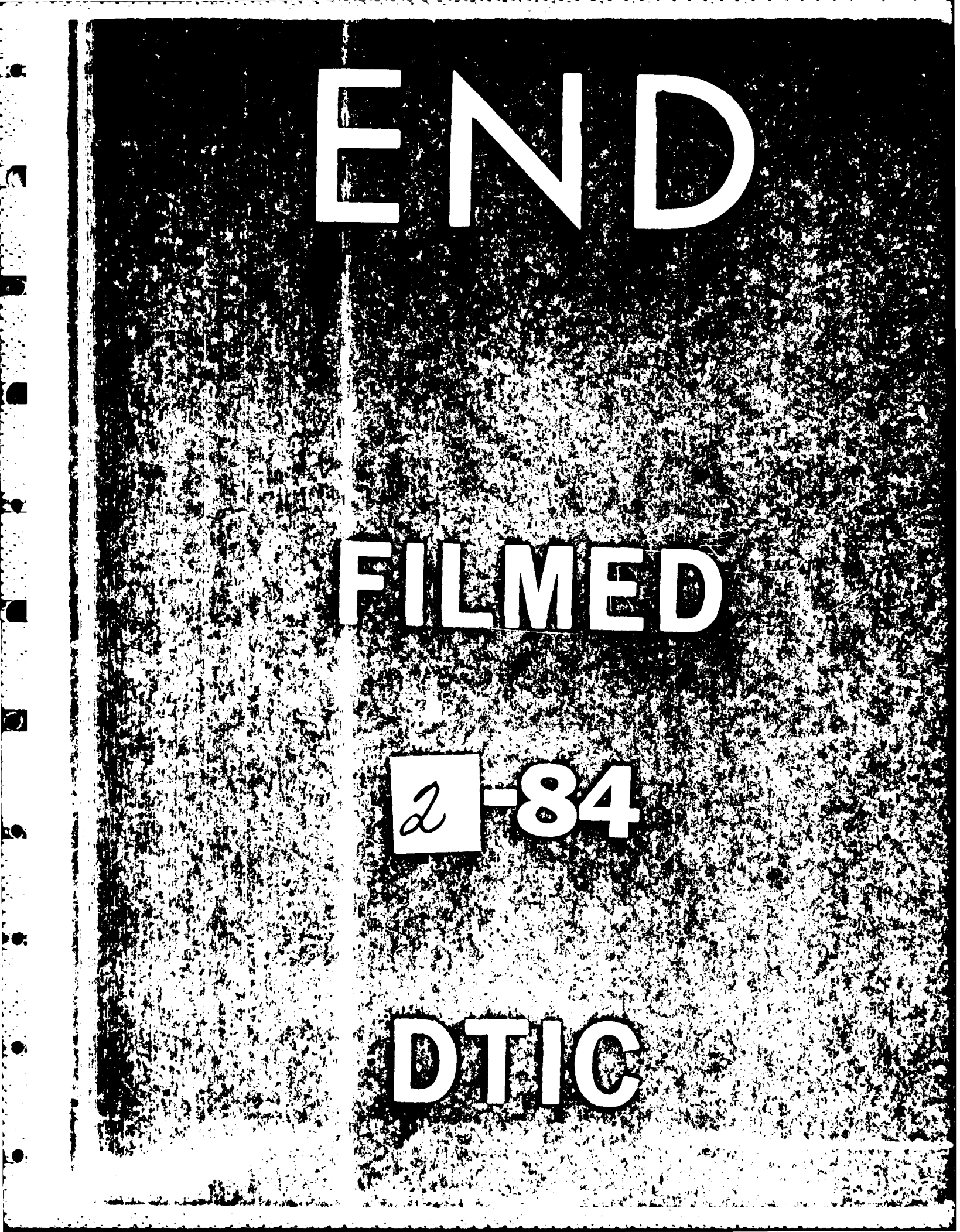
Siegel, A. I. Trainability testing for Navy selection and classification (NPRDC Tech. Rep. 83-28). San Diego: Navy Personnel Research and Development Center, July 1983. (AD-A131 124)

Siegel, A. I., & Wiesen, J. P. Experimental procedures for the classification of naval personnel (NPRDC Tech. Rep. 77-3). San Diego: Navy Personnel Research and Development Center, January 1977. (AD-A035 744)

Vineberg, R., & Joyner, J. N. Prediction of job performance: Review of military studies (NPRDC Tech. Rep. 82-37). San Diego: Navy Personnel Research and Development Center, March 1982. (AD-A113 208)

DISTRIBUTION LIST

Assistant Secretary of the Navy (Manpower & Reserve Affairs) (OASN(M&RA))
Deputy Assistant Secretary of the Navy (Manpower) (OASN(M&RA))
Deputy Assistant Secretary of the Navy/Military Personnel and Force Management, Director Accessions Policy
Chief of Naval Operations (OP-01), (OP-11), (OP-115) (2), (OP-12), (OP-13), (OP-135C4), (OP-14), (OP-140F2), (OP-15), (OP-987H)
Chief of Naval Material (NMAT 05), (NMAT 0722)
Director, Naval Civilian Personnel Command
Deputy Chief of Naval Material (Technology)
Chief of Naval Research (Code 200), (Code 270), (Code 440) (3), (Code 442), (Code 442PT)
Chief of Naval Education and Training (00A), (N-21)
Chief of Naval Technical Training (016)
Commandant of the Marine Corps (MPI-20)
Commander Fleet Training Group, Pearl Harbor
Commander Military Enlistment Processing Command
Commander Naval Military Personnel Command (NMPC-013C), (NMPC-48)
Commander Navy Recruiting Command
Commanding Officer, Naval Aerospace Medical Institute (Library Code 12) (2)
Commanding Officer, Naval Regional Medical Center, Medical Library
Commanding Officer, Service School Command (Code 3200)
Commanding Officer, Naval Technical Training Center, Corry Station (Code 101B)
Commanding Officer, Naval Training Equipment Center (Technical Library) (5) ,(Code N-1)
Commanding Officer, Office of Naval Research Branch Office, Chicago (Coordinator for Psychological Sciences)
President, Naval War College
Superintendent, Naval Postgraduate School
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (PERI-ASL), (PERI-ZT), (PERI-SZ)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Manpower and Personnel Division), (Scientific and Technical Information Office)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT)
Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/LR)
Commanding Officer, U.S. Coast Guard Research and Development Center, Avery Point
Superintendent, U.S. Coast Guard Academy
Institute for Defense Analyses, Science and Technology Division
Defense Technical Information Center (DDA) (12)


END
FILMED
2-84
DTIC