

DDC FILE COPY

MA069371

NPRDC TR 79-17

May 1979

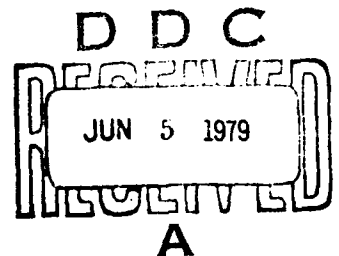
**AN INVENTORY BATTERY TO PREDICT NAVY AND MARINE CORPS
RECRUITER PERFORMANCE: DEVELOPMENT AND VALIDATION**

Walter C. Borman
Jody L. Toquam
Rodney L. Rosse

Personnel Decision Research Institute
Minneapolis, Minnesota 55402

Reviewed by
Martin F. Wiskoff

Approved by
James J. Ragan
Technical Director



Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPRDC/TR-79-17	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AN INVENTORY BATTERY TO PREDICT NAVY AND MARINE CORPS RECRUITER PERFORMANCE: DEVELOPMENT AND VALIDATION.		5. TYPE OF REPORT & PERIOD COVERED Interim Report May 1976 - August 1977
6. AUTHOR(s) W. C. Borman, J. L. Toquam R. L. Rosse		7. PERFORMING ORG. REPORT NUMBER
8. PERFORMING ORGANIZATION NAME AND ADDRESS Personnel Decisions Research Institute 821 Marquette Avenue, Foshay Tower Minneapolis, Minnesota 55402		9. CONTRACT OR GRANT NUMBER(s) N00123-76-C-1284
10. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 603707N ZPN01.06
12. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) ZPN01		13. REPORT DATE May 1979
		14. NUMBER OF PAGES 132
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
18. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
19. SUPPLEMENTARY NOTES Walter C. /Borman, Jody L. Toquam Rodney L. /Rosse		
20. KEY WORDS (Continue on reverse side if necessary and identify by block number) Test validation Navy recruiters Marine Corps recruiters Criterion measurement Validity study Criterion development		
21. ABSTRACT (Continue on reverse side if necessary and identify by block number) The objective of this study was to develop paper-and-pencil predictors of Navy and Marine recruiter performance and evaluate their validity. Accordingly, several measures of personality, vocational interests, and background were prepared (or selected) and administered to a geographically representative sample totaling 329 Navy and 118 Marine Corps recruiters. Scores on the predictor battery's items and scales were correlated with performance scores developed from supervisory, peer, and self ratings and from production data		

DD FORM 1473

JAN 73

EDITION OF 1 NOV 68 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

390 698

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

(i.e., numbers of recruits enlisted). Estimated cross-validities for predictor composites were significantly different from zero for four of the five performance criteria in the Navy sample. They ranged from .17 to .31. Corresponding validity estimates for the Marine Corps sample ranged from .22 to .38, (*less than* $p < .01$ for three criteria, $p < .05$ for two criteria).

Recommendations from the study included:

1. Examine the predictive validity of the predictor composites developed in this project,
2. Assess the potential fakability of the predictor composites, *and*
3. Develop additional paper-and-pencil measures of constructs that this study suggests are valid indicators of Navy and Marine Corps recruiter success.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

FOREWORD

This study constitutes the first phase in the development of techniques for identifying those individuals with the greatest likelihood of performing effectively as Navy and Marine Corps recruiters. It includes evaluations of existing selection procedures as well as experimental procedures.

Work was conducted in support of Advanced Development Subproject ZPN01.06 (Advanced Navy Recruiting System) and Exploratory Development Task Area ZF55.521.001.101 (Marine Corps Personnel Resources Management), under the joint sponsorship of Navy Recruiting Command and Headquarters, Marine Corps (MPI-20).

The results of the study will be used in a second phase that will evaluate the predictive validity and fakability of some of the predictor composites examined.

Appreciation is expressed to LCDR Hank Levein of the Navy Recruiting Command and to Major William Blaha, Headquarters, Marine Corps, for their skillful coordination efforts in providing the samples used in the study. Additionally, special thanks go to the participating Navy and Marine Corps recruiters who provided their time and expertise to aid us during the project.

The technical monitor was Dr. Norman Abrahams.

DONALD F. PARKER
Commanding Officer

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification	
By _____	
Distribution _____	
Availability Codes	
Dist	Availand/or special
A	

SUMMARY

Problem

The U.S. Navy and Marine Corps are very concerned about the number and quality of young persons brought into these services. Recruiters charged with the responsibility for contacting prospective recruits and selling them on a Navy or Marine Corps career are clearly the key personnel for successful recruitment of well-qualified men and women. Therefore, the selection of recruiters with good potential for success on this job becomes a vitally important problem.

Objective

The objective of this study was to develop pencil-and-paper predictors of Navy and Marine Corps recruiter effectiveness and evaluate the validity of these measures.

Approach

Based on a literature review of previous military recruiter selection studies, a trial predictor battery of personality, interest, and biographical items and scales was developed. The battery was administered to a geographically representative sample totalling 329 Navy and 118 Marine Corps recruiters. Scores on the predictor measures were correlated with performance ratings and with an objective effectiveness index in a concurrent validation design. Relationships between the various predictors and performance criteria were assessed.

Findings

1. Estimated cross-validities for predictor composites against four of the five performance criteria were significantly different from zero at the .01 level for the Navy sample. These estimated validity coefficients ranged from .17 to .31.
2. Comparable validity estimates for the Marine Corps sample ranged from .22 to .38, with all five predictor composite-performance criterion relationships significantly different from zero at the .05 level.

Conclusions

The predictor battery developed in this study shows promise for aiding Navy and Marine Corps decision-makers in selecting recruiters for the two services.

Recommendations

1. Examine the predictive validity of the composites developed in this research.
2. Examine the potential fakability of the selection instruments.
3. Develop additional paper-and-pencil measures of constructs that have proven to be valid predictors of recruiter performance.

CONTENTS

	Page
INTRODUCTION	1
Problem	1
Purpose	1
Background	1
Navy Studies	1
Marine Corps Studies	5
Army Studies	10
The Air Force Study	13
Implications of These Studies for the Present Project	16
METHOD	17
Summary of Methods	17
Pretest Procedures	17
Pretest Results	19
Interrater Agreement	19
Leniency	22
Restriction of Range	22
Halo	22
Determining Dimensionality of the Ratings	22
Developing an Objective Production Index	23
Pretest Performance Criteria	23
Using Pretest Data to Refine Rating Scale Materials	23
Using Pretest Responses to Revise the Trial Predictor Battery	24
Forming Predictor-Criterion Correlation Hypotheses Using Pretest Data	26
Assigning Cross-Validation Strategies to Different Sections of the Predictor Battery	27
Personality Scales	27
SCII	27
Biographical Survey	27
Description of the Sample for the Main Study	28
Procedures for the Main Study	28
RESULTS	31
Criterion Development Results	31
Interrater Agreement	31
Leniency	31
Restriction of Range	31
Halo	31
Determining Dimensionality of the Ratings	34
Developing the Objective Production Index	34

	Page
Performance Criteria	37
Relationships Between Criteria	37
Trial Predictor Battery Validation Results	37
Cross-Validation Results for Preselected Keys	38
Monte Carlo Cross-Validation Results	38
Estimated Validity for the Final Composites Pooled Together . .	43
Black-White Comparisons: An Examination of Bias and Differential Validity in the Navy Sample	43
CONCLUSIONS	47
Performance Measurement: Differences Between the Present Study and Previous Efforts	47
Relationships Between Predictor Composites and Recruiter Performance: Some Perspective	48
Comparison of 16PF and Predictor Composite Validity Results . . .	50
RECOMMENDATIONS	53
Studying the Predictive Validity of the Predictor Composite . . .	53
Studying the "Fakability" of the Predictor Battery	53
Developing Additional Measures of Valid Constructs	54
REFERENCES	55
APPENDIX A—16PF VALIDITY STUDY	A-0
APPENDIX B—SCALE DEVELOPMENT STRATEGIES AND DEFINITIONS OF RATING DIMENSIONS	B-0
APPENDIX C—MAIN SAMPLE PERFORMANCE RATING SCALE PACKAGE	C-0
APPENDIX D—NAVY RECRUITER BIOGRAPHICAL AND OPINION SURVEY	D-0
APPENDIX E—DESCRIPTION OF THE MONTE CARLO CROSS-VALIDATION PROCEDURE	E-0
APPENDIX F—HYPOTHESIZED PREDICTOR-PERFORMANCE CRITERION RELATIONSHIPS	F-0
APPENDIX G—IMPLEMENTATION OF THE PREDICTOR TEST BATTERY FOR USE IN SELECTING NAVY AND MARINE CORPS RECRUITERS	G-0
DISTRIBUTION LIST	

LIST OF TABLES

	Page
1. Measures Included in Pretest Predictor Battery	18
2. Pretest Rating Dimension: Results of Analyses	20
3. Intercorrelations and Reliabilities of the Final Five Performance Criteria—Pretest Data	24
4. Correlations Between Pretest Predictor Battery Personality Scales and the Five Pretest Performance Criteria	25
5. Interrater Agreement Indices for Self, Peer, and Supervisor Ratings	32
6. Means and Standard Deviations of Ratings Pooled Across the Self-Peer-Supervisor Rating Sources	33
7. Three-Factor Varimax Rotated Solution for Self-Peer-Supervisor Ratings Pooled Together: Navy and Marine Corps Main Sample Combined	35
8. Intercorrelations and Reliabilities of the Five Criteria . . .	36
9. Correlations Between Predictor Battery Personality Scales and the Five Performance Criteria: Main Sample, Navy	39
10. Correlations Between Predictor Battery Personality Scales and the Five Performance Criteria: Main Sample, Marine Corps	40
11. Cross-Validation Results for the Personality Composites	41
12. Cross-Validation Results for the SCII Keys	41
13. Monte Carlo Results—SCII, Navy Sample	42
14. Monte Carlo Results—Biographical Survey	42
15. Intercorrelations of Predictor Composites/Keys and Estimated Validity for the Final Composites Pooled Together . .	44
16. Black-White Differences: Means and Standard Deviations on Predictor Composites and Performance Criteria	45
17. Black-White Differences: Validity Coefficients for Final Predictor Composites	46

INTRODUCTION

Problem

The U.S. Navy and Marine Corps need highly qualified young people to fill a variety of jobs in the two services. Skilled recruiters are needed to contact qualified prospects and sell them on a Navy or Marine Corps enlistment. Although training and competent management can help create an effective recruiter force, selecting individuals with excellent potential for success as recruiters is of equal importance.

Purpose

The purpose of this study was to develop and evaluate a set of paper-and-pencil measures capable of predicting the performance of prospective recruiters. These measures can greatly aid the Navy and Marine Corps recruiter selection effort.

Background

The selection of effective recruiters responsible for enlisting qualified personnel has been of great concern to all branches of the Armed Forces. Selection batteries tapping a wide variety of individual differences (e.g., biographical data, occupational interest, personality, and personal preferences) have been examined to determine their usefulness in predicting military recruiter success. In addition, criteria designed to reflect recruiter effectiveness (e.g., supervisor ratings and production data) have been employed to evaluate the validity of the screening batteries. This section presents an overview of recruiter selection studies conducted by the Navy, Marine Corps, Army, and Air Force.

Navy Studies

Wollack and Kipnis (1960) developed a Navy recruiter selection battery and conducted a concurrent validity study to assess the battery's usefulness as a screening device. The predictor battery employed by the authors was composed of 13 measures designed to reflect fluency of expression, knowledge of the Navy, interest in recruiting activities, and general aptitude. The following instruments constituted the predictor battery:

1. Fluency of Expression: Four timed inventories designed to measure a subject's verbal abilities. (These timed inventories were administered to the cross-validation group only.)
 - a. Thing Listing Test--requires the subjects to write the names of as many things as possible.
 - b. First Letters Test--requires the subjects to write as many words as they can think of beginning with letter "b."
 - c. Four-Word Combination Test--requires the examinees to compose four-word sentences using the four letters presented in each item as the initial letters of the words.

d. Inventive Opposite Test—a word is presented along with the first letter of one of its synonyms and the first letter of an antonym. The subject is required to fill in the synonym and antonym for each word in the series.

2. Knowledge of and Enthusiasm for the Navy

a. Navy Knowledge—a 60-item inventory designed to measure a subject's knowledge of Navy history and traditions.

b. Career Preference Scale—30 items intended to measure a subject's attitude toward a Navy career.

c. Career Motivation Survey—a 39-item attitude scale, measuring attitudes toward various aspects of Navy life.

3. Vocational and Sports Interests

a. Kuder Preference Record—a published inventory designed to measure a subject's interest in 10 broad occupational areas.

b. Sports Inventory—a 50-item inventory dealing with rules and plays of various sports, designed to measure a subject's interest in athletic activities.

4. Aptitudes in four areas measured by the Basic Test Battery.

a. General Intelligence.

b. Arithmetic Reasoning.

c. Mechanical Comprehension.

d. Clerical Skills.

The validation sample consisted of 410 active recruiters representing 40 main recruiting stations, substations, and branch stations. The sample was formed by contacting the commanding officers from the 40 recruiting stations and asking each to nominate the most effective ($N = 205$) and least effective ($N = 205$) recruiters from their respective stations. The nominations were employed as the criterion measure against which the predictors were validated.

To validate the selection battery, it was administered to members of the sample by an officer from each recruiting station. The completed batteries were returned to the authors, and the data were analyzed.

The Navy Knowledge Test, Career Preference Scale, Career Motivation Survey, and Sports Inventory were item-analyzed using the effective and ineffective nominations as the criterion. Items that discriminated between effective and ineffective recruiters beyond the .20 level of confidence were retained for cross-validation. Seven items from the Navy Knowledge Test discriminated beyond the .20 level, as did 15 items from the Career

Preference Scale, 34 items from the Career Motivation Survey, 13 items from the Sports Inventory, and four scales from the Kuder.

Another variable--Navy service rating classification (duty prior to serving as a recruiter)--was analyzed to determine whether any differences existed between the effective and ineffective groups. Some significant differences were found. The effective recruiter group contained larger proportions of men in Deck ($p < .01$) and Aviation ratings ($p < .01$). The ineffective group contained significantly greater proportions of men in Engineering and Hull ratings ($p < .01$) and in Construction ratings ($p < .02$). These significant rating classification differences were retained for subsequent cross-validation.

The sample used to cross-validate the predictor battery consisted of 260 students attending a 6-week recruiter course. The predictor battery was administered to the cross-validation group while they were attending school. Approximately 1 year after the initial testing, when all students had been assigned to recruiting duty, ratings were collected from the two supervisors most familiar with each recruiter. The evaluation forms contained four 14-point scales: Technical Competence, Willingness to Work, Military Manner, and Adaptability. The forms also required the supervisors to answer the following two questions using 5-point scales:

1. If you had been given the chance, would you have recommended the man for recruiting duty?
2. Is this man effective in recruiting personnel?

The six ratings were treated as six different criterion measures. Recruiter scores on each dimension were computed by adding the two supervisory ratings together. When only a single supervisory rating was available, the scores were simply doubled. A total of 222 recruiters were rated by two supervisors. Interrater reliabilities ranged from .56 to .66, indicating an acceptable level of rater agreement.

Product-moment correlations between the predictors selected previously and the criterion measures were computed using the cross-validation sample of 260. Three of the variables significantly predicted the overall evaluation of recruiting effectiveness: the Kuder Persuasive Scale ($r = .24$, $p < .01$), the Kuder Scientific Scale ($r = -.17$, $p < .01$), and the Career Motivation Survey item analysis key ($r = .13$, $p < .05$). The Kuder Persuasive Scale also correlated significantly ($p < .01$) with each other criterion (except Adaptability). In addition, the Kuder Scientific Scale correlated significantly ($p < .01$) in the negative direction with the rating dimensions, Willingness to Work and Recommendation for Recruiting Duty, and the Career Motivation Survey Key correlated significantly ($p < .01$) with Technical Competence. None of the other predictors correlated significantly with any of the criteria with the exception of one fluency-of-expression measure, the First Letters Test, which yielded a correlation of $-.18$ ($p < .01$) with Military Manner. A multiple correlation was computed using the Persuasive, Scientific, and Career Motivation Survey predictors against the criterion measure of overall recruiter effectiveness. The result ($R = .26$) was not significantly greater than the correlation between that performance dimension and the Persuasive Scale alone.

Thus, results of the cross-validation analysis indicate that an effective Navy recruiter has persuasive interests, is not overly interested in scholarly pursuits, and believes in the value of a Navy career.

The Wollack and Kipnis (1960) study employed a predictor battery containing a variety of tests and inventories. Although the reliability of the criterion nominations was reasonably high, only a limited number of these scales and item keys cross-validated significantly. This may have occurred because the raters were making their evaluations on the basis of reputation instead of performance or because many of the individual differences that are truly predictive of recruiter success simply were not included in the battery.

Krug (1972) administered an experimental inventory, the 16PF personality inventory, to Navy enlisted recruiters and officer recruiters to determine the usefulness of the instrument for predicting recruiter effectiveness.

The inventory, which was administered to recruiters in 41 recruiting stations across the continental United States, consisted of the following items:

1. The 16PF questionnaire, Form A, 1967 edition--a personality test designed to measure an individual's personality in terms of 16 basic factors.
2. A 25-item experimental supplement including 7 items designed to measure motivational distortion (a lie scale) and 18 items designed to measure strength of motivation to succeed as a recruiter.
3. Seven biographical items yielding information about years of service, age, sex, marital status, number of dependents, years of formal education, and population of the subject's home of record.

Commanding officers from each of the 41 Navy recruiting stations were asked to nominate five enlisted recruiters and one officer recruiter from among the top 50 percent and five enlisted recruiters and one officer recruiter from among the bottom 50 percent of recruiters presently on duty. These nominations of recruiter effectiveness served as the criterion against which the 16PF, supplemental, and demographic variables were validated.

A stepwise multiple regression of the several predictor variables on the criterion for the enlisted sample resulted in a multiple correlation coefficient of .40, significant beyond the .01 level. When the same equation was applied to the officer recruiter cross-validation sample ($N = 74$), a multiple correlation coefficient of .25 was obtained ($p < .05$).

According to these results, the effective Navy recruiter is typically married, has more years of formal education, tends to be warm and outgoing, dominant, aggressive, self-assured, and relatively conservative politically.

Although the results from the cross-validation sample supported the validity of this instrument, it is important to take a closer look at the regression equation. The most influential variable, Marital Status, was assigned a high positive weight in the prediction equation. Yet, only 13 of the 383 enlisted recruiters in the sample were unmarried. In a selection

situation, a low base-rate of unmarried recruiters would make that variable less useful than a variable with a more balanced base-rate.

Abrahams, Neumann, and Rimland (1973) employed the Strong Vocational Interest Blank (SVIB) to develop an empirically keyed Recruiter Interest Scale (RIS) for use in selecting Navy recruiters. The first step in deriving the key was to request that the Commanding Officer in each of the 42 main recruiting districts identify the five least effective and five most effective recruiters. These CO nominations were used as the criterion to develop and validate RIS. Thus, ten SVIB inventories were mailed to each of the 42 districts to be administered to the recruiters selected for the study.

Complete information was obtained from 356 recruiters representing 36 of the recruiting districts. One-half of the sample ($N = 178$) was used to develop the RIS, while the other half ($N = 178$) was used as the cross-validation group. Each of the 399 SVIB items was weighted according to the proportion of most- and least-effective recruiters endorsing the response. If a response was endorsed more often by the most effective recruiters it was weighted +1. Conversely, if a response was endorsed more often by the least effective recruiters it was weighted -1. The 115 items with the largest endorsement differences were included in the SVIB-RIS-1.

The "holdout" group was used to cross-validate the SVIB-RIS-1. Responses to the SVIB items were assigned weights according to the differential endorsement patterns of the previous group. The scores for each recruiter were tabulated, arranged from high to low (high scores indicated potentially effective recruiter performance and low scores indicated potentially ineffective performance) and the cross-validation sample was then divided into quartiles according to RIS scores. The top quartile (highest RIS scores) contained three times as many effective recruiters as did the bottom quartile; conversely, the bottom quartile contained three times as many ineffective recruiters as did the top RIS group. Clearly, the cross-validity obtained for the RIS-1 was impressive. Accordingly, the authors recommended using the SVIB-RIS-1 as a selection device to exclude low-scoring individuals from recruiter service. Abrahams et al. suggested that a cutoff score would largely depend on the number of applicants and the number of recruiters needed. The authors also recommended using supplemental inventories (e.g., biographical items found effective in predicting sales performance) to increase performance predictability.

In addition, they recommended conducting additional research with the SVIB to provide greater stability in establishing the scoring weights and in assessing validity. Abrahams et al. stress that "a better criterion of recruiter effectiveness must be developed" and that additional factors of recruiter performance should be taken into account in future validity studies.

Marine Corps Studies

Bennett and Habers (1973) assessed the degree to which individual recruiter characteristics, geographic assignment, and "evaluation" variables influence Marine Corps recruiter productivity. The authors collected data from 259 Marine Corps recruiters representing 29 recruiting stations throughout the United States. The information was collected from performance records at the Marine Corps Recruiter Depot, the enlisted Master File, and a survey of

recruiters and their jobs conducted by the Office of Manpower Utilization. The variables were divided into three categories:

1. Selection variables--General Classification Test scores, age, race, level of education, number of dependents, previous service as a career planner or drill instructor, method of assignment to recruiting duty (volunteer or assigned), and opinion about recruiting duty as a financial hardship.

2. Deployment variables--whether the recruiters are assigned in their home states; distance from home state; type of area assigned to--urban, suburban, or rural; number of times assigned; hours per week spent on recruiting; and percentage of time spent out of the office recruiting.

3. Evaluation variables--number of months the individual has been on his current tour of duty and the recruiter's percentile rank in Marine Corps Recruiter Class.

Gross productivity (i.e., the average number of recruits enlisted per month) was employed as the measure of performance for each recruiter. However, Bennett and Habers noted that differences in gross productivity cannot be completely attributed to differences in the performance of individual recruiters, but may be due in part to a variety of regional differences (e.g., the number of qualified prospects and local attitudes toward military service). For this reason, the sample was divided into two groups on the basis of the average regional enlistment rates for 1971. The first group contained those recruiters from high enlistment recruiting stations ($N = 109$) and the second group contained recruiters serving in the low enlistment areas ($N = 150$).

The authors then used multiple regression to determine the magnitude of the relationship between each variable and recruiter productivity. A separate regression equation was developed for each enlistment area group. In the high enlistment area group, three variables were found to be significantly related to productivity. The first variable, type of area, showed that urban and suburban recruiters enlisted more persons per month ($p < .01$) than rural recruiters. The second variable, geographical assignment, showed that recruiters stationed in their home state enlisted more persons per month ($p < .01$) than those stationed more than 500 miles outside their home state. The third indicated that productivity was negatively related to increased reassignments ($p < .05$). These three variables accounted for 35 percent of the productivity variance.

In the low enlistment areas, two variables correlated significantly with productivity. The hardship variable indicated that those recruiters who feel recruiting duty is a financial hardship enlist more persons per month ($p < .05$) than those who do not. The second significant variable, service previously as a career planner, indicates that recruiters with prior service enlisted a greater number of persons per month ($p < .05$) than did those who had no experience as career planners. The two variables account for only 12 percent of the productivity variance, however.

It should be noted that the regression equations obtained in the study were not cross-validated, and thus the relationships found here may well overestimate the true usefulness of these variables.

Larriva (1975) applied the Navy's recruiter selection battery, the 16PF-M, to a Marine Corps sample in a concurrent validity study.

The inventory, which was administered to all active recruiters from the 9th Marine Corps District ($N = 470$), consisted of the same items used in the Krug (1972) study--the 16PF questionnaire, a 25-item supplement, and 7 biographical items.

The 16PF-M inventories were mailed to the 9th Marine Corps Recruiting command, which arranged for their administration to the recruiters. The Recruiting Command also made arrangements to record the number of Non-Prior Service (NPS) accessions (i.e., recruits enlisted) for each recruiter throughout the year 1974. The accessions data, which were employed as the criterion measure, were provided to the author quarterly.

After all the completed 16PF-M inventories were returned and the first-quarter accession data were reported, Larriva analyzed the data using the formula developed by the Navy to predict recruiter success. This analysis indicated that the Navy formula was not a valid equation for predicting Marine Corps recruiter production.

Larriva conjectured that the accessions criterion might not provide relevant performance information; thus, he generated several performance indices on the basis of assumptions about the recruiting environment and geographical differences in production. Using first- and second-quarter accession data, each performance index was employed as the criterion measure and correlated with the personality, demographic, and experimental variables of the 16PF-M. Larriva then examined the resulting predictor-criterion relationships and selected the criterion index that yielded the most valid multiple correlation coefficient. The index selected, referred to as Pl_6 , indicated the number of accessions of urban and rural recruiters separately and corrected for geographical differences in relative performance of recruiters. Accordingly, the validation sample was split into an Urban Recruiter Group ($N = 308$) and a Rural Recruiter Group ($N = 162$) and a weight correcting for geographical differences was applied to production scores. The data from each group were analyzed separately.

Larriva also chose to generate samples of urban and rural recruiters that contained only the best and poorest recruiters. The reduced samples were referred to as the Urban Hi-Low ($N = 122$) and the Rural Hi-Low ($N = 96$) groups.

The Pl_6 criterion was then used as the performance measure, and step-wise multiple linear regression was utilized to generate four formulas (one for each of the samples--Urban, Urban Hi-Low, Rural, and Rural Hi-Low) for predicting recruiter success from the 16PF-M data.

The regression technique yielded prediction equations and multiple correlation coefficients for each sample--Urban $R = .43$ ($p < .01$), Urban Hi-Low $R = .59$ ($p < .01$), Rural $R = .33$ ($p < .15$), and Rural Hi-Low $R = .44$ ($p < .42$).

The means and standard deviations of the predicted and actual scores were calculated for the four data bases and were used to establish cutoff scores. These cutoff scores were then used hypothetically to assess the effectiveness of the prediction equations to select or delete recruiters in a cross-validation sample.

The cross-validation sample consisted of 98 recruiters who were not included in the validation sample but who had a minimum of three months accession data. As prescribed by the production index, the sample was split into an Urban Recruiter Group ($N = 66$) and a Rural Recruiter Group ($N = 32$), and effectiveness scores (the Pl_6) were generated for each recruiter.

The urban prediction equations were applied to the urban cross-validation sample. The results of the prediction equations were compared according to accuracy and expected increase in productivity. The same procedures were followed to compare the results of the Rural prediction equation and Rural Hi-Low prediction equation when applied to the rural cross-validation sample. Results for the two comparisons are listed below:

	<u>Urban</u>	<u>Urban Hi-Low</u>	<u>Rural</u>	<u>Rural Hi-Low</u>
Number deleted by cutting score	24	34	12	12
Number below mean deleted	21	26	8	10
Percent correct deletion	88	76	75	83
Mean production score of those deleted	9.29	9.91	10.06	9.02
Mean production score of those remaining	12.29	12.54	12.39	13.02
Expected percentage increase in productivity	10	12	7.6	13

From these comparisons, Larriva concluded that the Urban predictor equation and Rural Hi-Low equation would prove successful in selecting Marine Corps recruiters. The author also suggests that both equations could be used to analyze each 16PF-M inventory, thus allowing the Marine Corps Recruiting Command to select the optimal setting, urban or rural, for each new recruiter.

In summary, the Navy prediction equation did not effectively predict Marine Corps recruiter success. However, when recruiting setting (urban-rural) and geographical differences were accounted for in the criterion, relatively successful prediction equations were formed, one for urban recruiters and one for rural recruiters. Cross-validity results suggest that the 16PF-M may be a useful selection device for screening candidates for the Marine Corps recruiter job. It would be of interest to apply the Navy

prediction equation for the 16PF-M to the cross-validation sample using the Pl_6 criterion. Apparently, Larriva never correlated the Navy prediction equation scores with the scores on the Pl_6 criterion.

The selection of a criterion in this research is of interest. Larriva essentially selected the criterion that correlated most highly with the predictors. A clearly more acceptable (and justifiable) method is to define the criterion, or criteria, as precisely as possible and then to select a measure that provides relevant and reliable measurement of each criterion without regard to the predictors. The cross-validation procedure employed in Larriva's study makes less severe the criticism just leveled, but the Pl_6 criterion would have been more justified as a measure of effectiveness if predictor data had not dictated its choice.

Graf and Brower (1976) investigated the usefulness of the Strong Vocational Interest Blank for selecting Marine Corps recruiters. SVIB inventories were administered to 98 recruiters working in the Los Angeles and Santa Ana areas. Of those 98, 77 recruiters completed and returned the SVIB inventories. Officers in the Los Angeles and Santa Ana recruiting stations were asked to provide effectiveness ratings for each of the recruiters. These ratings consisted of a 3-point scale representing below-average, average, and above-average recruiter performance, and were employed as the criterion measure.

The authors recognized that the Navy Recruiter Interest Scale developed by Abrahams et al. (1973) on a sample of Navy recruiters might predict Marine Corps recruiter effectiveness as well. Thus, Marine Corps recruiter SVIB responses were scored using the RIS scale and its validity was evaluated. First, a t-test was used to compare the rated performance of recruiters from the upper and lower 35 percent of the RIS score distribution. The results were in the predicted direction (significant beyond the .05 level); the higher the RIS score, the greater the likelihood that the recruiter was in the effective group.

Second, recruiters were dichotomized as above average or below average according to the supervisory field ratings, and the two groups' RIS scores were compared. A t-test once again showed a significant difference (beyond the .05 level) in the predicted direction (r_{pt} , biserial = .30, $p < .05$).

Although the RIS effectively discriminated between above-average and below-average Marine Corps recruiters, Graf and Brower hypothesized that differences between Navy and Marine Corps recruiters might warrant a separate selection device. The authors elected to develop a Marine Corps Recruiter Interest Scale (MCRIS) key based on Marine Corps recruiter responses. To permit development of a separate key, the SVIB was administered to a second sample of marine Corps recruiters ($N = 91$) attending the Marine Corps recruiting school. Two years after the SVIB administration, the students were located to collect recruiter-effectiveness ratings from their respective supervisors.

The criterion measure of recruiter effectiveness was used to dichotomize the combined sample ($N = 168$, consisting of 77 recruiters from the

original sample and 91 Marine Corps students at the recruiter school) into below-average and above-average recruiters based on the supervisor ratings. The 339 SVIB items were then weighted according to the differential endorsement patterns of the two groups. Items were included in the MCRIS if endorsement differences between above- and below-average recruiters exceeded 15 percent. This procedure resulted in 75 items for the key. (Interestingly, only 13 of the 75 items in MCRIS are also found in the RIS.)

The combined Marine Corps recruiter sample was scored using the RIS and the MCRIS. The validity coefficients were .67 for the MCRIS and .29 for the RIS. However, these two coefficients cannot be compared directly because the RIS validity coefficient reflects the key's cross-validity, while the cross-validity of the MCRIS cannot be gauged from these data.

Graf and Brower recommend assessing the MCRIS's utility as a selection device by evaluating the cross-validity of the key on a larger, more representative sample of Marine Corps recruiters. The authors also indicate concern for their criterion measure of recruiter effectiveness. They recommend attempting to develop a more reliable method of measuring recruiter performance.

Army Studies

Brogden and Taylor (1949) reexamined the results of their earlier study designed to predict Army recruiter effectiveness and evaluated the validity of the test battery using a different criterion measure of recruiter effectiveness. The initial study had employed four paper-and-pencil predictors of recruiter performance, including a measure of vocational interests and three measures tapping interests, hobbies, and backgrounds. The four predictor scales had been item-analyzed against a production criterion—the number of recruits obtained by each recruiter per hour on recruiting duty. The empirically keyed predictors were cross-validated using a sample of 475 recruiters, yielding a validity coefficient of .18.

Brogden and Taylor postulated that the low validity found in the initial study was due to the unreliability of the production criterion and to the contamination potentially present in such global objective indices. The authors elected to use a new criterion: a dichotomy of retention vs. nonretention on recruiting duty. The predictor data collected in the initial study was used to evaluate the validity of that information for predicting recruiter turnover. When the turnover criterion was applied to the summed scores of the four inventories, it yielded a biserial validity coefficient of .36.

The focus of the Brogden and Taylor study was on the criterion used to validate predictors used for selecting Army recruiters. The initial production criterion was dismissed because it was thought to contain many factors not associated with recruiter effectiveness. The authors recognized that the turnover criterion also contained some extraneous factors unrelated to recruiter effectiveness (e.g., completing tour of duty, requesting a transfer). When the new criterion was correlated with the predictors, however, a higher validity was obtained. This increase in validity may be due in part to the relatively low reliability and significant contamination

associated with the production criterion. It is also possible, of course, that measures such as those used here are simply better predictors of turnover than they are of performance.

Brown, Wood, and Harris (1975) sought to develop a valid criterion measure of recruiter effectiveness and to develop an improved Army recruiter selection instrument. In the first phase of their project, the authors focused on a criterion measure used unsuccessfully in previous recruiter studies--total production scores. Brown et al. noted that these figures were confounded by "opportunity bias" arising from territorial factors that fall outside the control of the individual recruiter. They hypothesized that if territorial factors were somehow partialled out of the production score variance, the result would be a more realistic, unbiased picture of recruiter effectiveness.

The authors began by identifying territorial variables that could bias production figures (e.g., the average number of enlistments per recruiter in recruiter's District Recruiting Command or DRC; the proportion of territory that is metropolitan, suburban, or rural; and the recruiters' experience). A total of 12 territorial factors were identified as potential biasing agents.

The next step was to obtain a nationwide random sample of Army recruiters. The sample was acquired by requesting each commander from the five Regional Recruiting Commands (RRCs) to supply names of 100 recruiters from their respective RRCs. Commanders selected recruiters randomly according to their social security numbers. In addition to the names, commanders were asked to supply total production figures for each recruiter for the period from July to December 1973. The territorial information of each DRC represented by recruiters in the sample was compiled by Army personnel from available records and from a special market survey they conducted.

The 12 factors were subjected to a stepwise multiple regression with total production scores. Three variables contributed significantly ($p < .05$) to the prediction of production scores--average production per recruiter in subject's DRC; proportion of all enlistees who chose the Army; and the proportion of the zone that is suburban. These three variables accounted for 51 percent of the production score variance.

The three territorial factors were used to develop a multiple regression equation that yielded predicted production scores for each recruiter. The predicted scores were employed to compute unbiased production scores. This was accomplished by computing the ratio of total production to predicted production scores multiplied by 100, and resulted in what the authors termed Benchmark Achievement Scores (BAS).

The BAS scores thus consisted of production data corrected for DRC production, proportion of young persons joining the Army in the DRC, and the proportion of the DRC that is suburban. The adjusted scores should provide relatively unbiased production information, an index of productivity that allows for differential opportunity to recruit successfully because of location.

The BAS scores were correlated with another production measure of recruiter effectiveness, Simple Achievement Scores. These scores were derived by computing a ratio of total production to the average production in the DRC (the variable that accounted for most of the variance in the multiple regression equation). Brown et al. recommended that the Simple Achievement Scores be used as future criterion measures because they correlate highly with the BAS ($r = .96$) and are easier to compute.

In the second phase of their research, Brown, Wood, and Harris developed an Army recruiter selection battery. Inventories were selected for the battery based on recruiter job behaviors and personal characteristics. This information was gained from interviews with 79 recruiters with high, average, and low records of success. Personnel from Army Recruiting Headquarters were also consulted about traits necessary for recruiter success.

The selection battery consisted of 15 paper-and-pencil inventories and one verbal performance test. Below is a list of the measures included in the battery.

1. Verbal Fluency--Recruiters were asked to simulate a presentation of the benefits of Army life to a prospective enlistee. Each presentation was recorded and scored by computing the ratio of the number of words spoken during the first 2 minutes of the presentation to the number of "ahs" spoken during the same length of time.
2. Sociability Measures--Four inventories were used to measure a recruiter's sociability and affiliative tendency.
3. Achievement Motivation--Three inventories were employed to tap the tendency to work hard to achieve self-appointed goals.
4. Empathy Measures--Four instruments were used to measure the ability to understand the point of view of others and the drive to win or complete a sale.
5. Rejection Tolerance Measure--One paper-and-pencil inventory was employed to measure a recruiter's tolerance to rejection, rebuffs, and insults.
6. Responsibility and Maturity Measures--Three instruments were used to tap information about a recruiter's ability to manage his personal, financial, and official duties.

When it became necessary to select subjects for the battery administration, the criterion development was not yet completed. Consequently, Brown et al. devised a Composite Supervisory Rating procedure to select two extreme groups of recruiters--the highly successful and the very unsuccessful. Two DRCs (the highest ranking DRC and the lowest ranking DRC, based on objectives achieved) were selected from each of the five RRCs. Five supervisors from each DRC were asked to nominate the ten best and ten poorest recruiters within their DRC. The five recruiters mentioned most often as the best recruiters were used in the High Criterion Group, while the five recruiters mentioned most often as the poorest recruiters were placed in the Low Criterion

Group ($N = 45$ for the High Criterion Group, and $N = 43$ for the Low Criterion Group). The 4- to 6-hour battery was administered to the High and Low Criterion Groups in each DRC, and additional information concerning each recruiter's race, religion, and aptitude scores was collected from Army personnel files. Means were computed for the keyed inventories and the aptitude scores for the High and Low Criterion Groups separately. None of the personality measures or aptitude scores differentiated significantly between the two groups; however, 20 background items did differentiate significantly ($p < .10$) between them. These items pertain to work habits, styles of handling finances and debts, educational background, and reaction to challenging or stressful situations. The verbal performance test also differentiated significantly ($p < .10$) between the two groups. Brown et al. suggest that because recruiters are a relatively homogeneous group required to meet several minimum qualifications (e.g., GCT scores, age, and rank) and because of the length of time in the service (in this study, the mean service time was 14 years), the recruiters may have formed similar attitudes and opinions, thereby limiting the variance in attitude, personal preference, and personality inventory scores. Conceivably, however, the low relationships between predictors and the criterion is a result of the criterion measure used. Utilizing the relatively unbiased production index, the Simple Achievement Score, might have led to higher validities.

The Air Force Study

The Air Force Recruiting Service recognized the need for a recruiter selection device and in 1966 enlisted the service of Massey and Mullins to develop such a selection instrument. The authors began by constructing an eight-inventory battery designed to measure qualities such as empathy, surgency (friendliness and sociability), and perseverance, all thought to be desirable in recruiters. The measures are listed below.

1. Airman Preference Test—a 43-item five-choice test designed to measure empathy. The five choices are to be ranked according to the average basic airman's preferences.
2. CEI Opinion Questionnaire—83 items describing a person's behavior, an individual's opinions, old proverbs, and expressions of interest. The examinee is to show agreement, disagreement, or no opinion. Three scores (Complexity, Ego Strength, and Introversion) were considered likely predictors of recruiter success.
3. Community Information—43 items to gain factual information about areas of former residence. This was designed on the rationale that a person who knows more about the community is one who is more interested in it and consequently one who should be a more effective recruiter.
4. Descriptive Adjective Inventory—182 paired adjectives used to describe a person's general appearance, usual behavior, and relations with others. The examinee selects the one of the pair that best describes him. Participants were scored on three scales (Surgency, Cooperativeness, and Orderliness).

5. FCSRI-A—190 items, pairs of statements, that refer to a person's appearance, usual behavior, and relations to others. Three scores (Surgency, Cooperativeness, and Orderliness, the same three variables as measured by the Descriptive Adjective Inventory) were considered likely predictors of recruiter effectiveness.

6. Recruiter Language—made up of four subtests:

a. 15 five-choice items—examinee must select the initial letter of a missing word. The definition is provided.

b. 15 five-choice items—examinees must select the correct arrangement of words appearing in progressive order. The words express varying degrees of the same general concept.

c. 10 five-choice items—examinee must select the best sentence with respect to grammar and usage.

d. 15 five-choice items—a sentence is provided; an examinee must select the interpretation that means most nearly the same as the sentence.

7. Texas Social Insight—40-item test to describe ordinary, unusual, or embarrassing situations. The examinees are to select the alternatives they see as the correct courses of action.

8. Word Power—test of ability to call to mind certain words. At a given signal the examinee records as many words as possible beginning and ending with a given letter of the alphabet.

In a preliminary validation step, the battery was administered to 210 recruiter students prior to entry in the recruiter course. Twenty-four scores, derived from the eight inventories, were compared to the criterion measure of graduation or elimination from the recruiter course (178 graduated and 32 were eliminated). Four of the scores—Recruiter Language, Community Information, FCSRI-A (Surgency), and the Texas Social Insight test—yielded significant point biserial correlations ($p < .01$).

A combination of the Recruiter Language, Community Information, and FCSRI-A (Surgency) scores yielded a multiple correlation of .23 ($p < .01$). (The Texas Social Insight Test did not add significantly to the multiple correlation.) The three inventories, with scoring weights of 2, 1, and 1 respectively, were employed as an interim selection battery, pending further validation of the eight-inventory composite.

The full eight-inventory battery was administered to a larger sample of recruiter students ($N = 1067$) prior to entry in the recruiter course. Background information including age, education, marital status, number of dependents, total active federal military service, and AQE scores was also tapped. Pass/fail information was subsequently obtained for each recruiter student.

The validation sample ($N = 965$, students with complete information) was randomly split into two groups--sample B ($N = 485$), the computing sample; and sample C ($N = 480$), the cross-validation group. Twenty-four scores, derived from the eight inventories, were generated for sample B recruiters. These scores, along with the background variables, were compared to the recruiter course pass/fail criterion. Various combinations of scores and background variables were also compared to the criterion measure. No combination of predictors significantly ($< .05$) increased prediction obtained using the three scores from the interim battery. The scoring weights of the interim battery (RSST-63) were altered a number of times and used to score sample B recruiters. These scores were validated against the pass/fail criterion. No alteration of regression weights surpassed the prediction capabilities of the interim battery with the original weights ($R = .34$, $p < .01$).

In addition, the regression weights of the predictor scores of sample B (the same scores and weights used in the interim battery) developed using the pass/fail criterion were applied to sample C responses. The scores of sample C recruiters were calculated and validated against the pass/fail criterion. This yielded a point biserial correlation of .21 ($p < .05$).

The same procedures were followed in validating the predictor scores against a second criterion--supervisor field ratings. These ratings were obtained from supervisors 1 year after each recruiter was placed on duty. Results showed that no combination of variables of the entire set significantly predicted the field ratings at the .05 level.

The authors also investigated the capability of the primary Air Force Specialty Code (AFSC) variable (service specialty from which a recruiter was drawn) to predict recruiter success. Recruiter training personnel postulated that those specialties that require much interpersonal contact would produce better recruiters than those specialties that require little contact. The primary AFSC variable was divided into four service specialty variables reflecting various degrees of required interpersonal contact. Additionally, 12 interaction variables were generated by multiplying each of the four primary service variables by aptitude, surgency, and dependability scores, thus creating 12 hybrid individual differences-AFSC variables. These 16 variables were then correlated with the field-rating criterion measure for a sample of 859 recruiter course graduates. Only one of the 16 correlated as highly as .07 ($p < .05$). The authors therefore discarded the primary AFSC as a predictor of recruiter success.

A second variable regarded as a possible predictor of field recruiter success was recruiter school advisor ratings. These ratings, collected while the recruiters were in recruiter school, consisted of a single 7-point scale ranging from outstanding to unsatisfactory. The advisor ratings were related to field ratings, yielding a correlation of .19.

The results of this study indicate that a battery consisting of the Recruiter Language Test, Community Information Inventory, and the FCSRI-A (Surgency) can be used to predict recruiter school success, but that neither this battery nor recruiter school advisor ratings can be used

to predict field recruiter success. The authors attribute the failure to predict field recruiter effectiveness to the supervisory performance rating criterion measure. They suggest that this kind of rating is inherently contaminated by several rater errors such as halo and leniency. Another difficulty may be that the battery the authors developed is simply more appropriate for predicting recruiter school success than field recruiter effectiveness. At any rate, Massey and Mullins believe that a more reliable and valid measure of recruiter effectiveness should be developed to provide a fair assessment of the utility of any paper-and-pencil measures purporting to predict recruiter performance.

Implications of These Studies for the Present Project

One difficulty that pervaded virtually every study reviewed is the "criterion problem": How can recruiter performance be measured in a reliable and valid manner? Many of the disappointing results obtained in these studies may in fact be linked to poor criterion measurement. Thus, it seems critical to measure recruiter performance very carefully in the present research.

Results of the studies discussed in this section also provide clues about predictors that might be used successfully as indicators of future effectiveness as a recruiter. First, the work of Abrahams, Neumann, and Rimland (1973) suggests that vocational interests may successfully predict military recruiter effectiveness. Results from the Wallack and Kipnis (1960) and Graf and Brower (1976) studies also indicate that interest inventories show promise as indicators of recruiter effectiveness.

The Krug (1972), Larriva (1975), and Brown, Wood, and Harris (1975) studies employed with some success background or biographical variables to predict recruiter performance. Although there is undoubtedly no single background ideal for recruiter success, it is possible that persons with certain kinds of backgrounds and experiences are more likely than others to perform effectively as military recruiters.

Personality or trait measures were also considered useful as predictors of recruiter performance. The Krug (1972) and Larriva (1975) studies found reasonably high relationships between personality scales and recruiter effectiveness.

On the negative side, abilities appear unrelated to recruiter performance. Intelligence and other kinds of ability measures generally failed to predict recruiter success.

Based in part on this literature review of test validation studies related to the military recruiter job, several personality, interest, and biographical items and scales were selected for the trial predictor battery. Additional predictors were selected based upon past experience in other selection studies and "clinical hunches" as to what kind of person makes a good Navy or Marine Corps recruiter. These "hunches" were based primarily on knowledge of the Navy and Marine Corps recruiter performance requirements gained from a previous study (Borman, Hough, & Dunnette, 1976). The predictors selected and their application in this study are described in the next section.

METHOD

Summary of Methods

A concurrent validity research design was employed in an attempt to identify paper-and-pencil predictors of Navy and Marine Corps recruiter effectiveness. Performance criteria for Navy and Marine Corps recruiters were derived from supervisor, peer, and self ratings on performance scales designed to evaluate military recruiters (Borman, Hough, & Dunnette, 1976), and on additional scales developed in this study. Also employed as a criterion was an objective measure of productivity that indexed the number of persons each recruiter brought into the Navy or Marine Corps over a 6-month period. This index was adjusted for geographical area.

Paper-and-pencil predictors that appeared to provide a valid indication of future success as a Navy or Marine Corps recruiter were selected after a review of earlier validation studies. A list of the predictors selected appears in Table 1. The predictor battery was administered to a nationwide sample totaling 329 Navy and 118 Marine Corps recruiters. Performance ratings and the productivity index were gathered for the same persons and relationships between predictors and performance indices were evaluated.

Also, 16PF (Cattell, 1974) scores were available for 191 of the Navy personnel in the sample who had completed that inventory before being selected for recruiting duty or during training for the job. The validity of the inventory was evaluated by correlating 16PF scores with the various performance criteria as described in Appendix A.

The methods related to the development and validation of the predictor battery are detailed below.

Pretest Procedures

Before the trial predictor battery and performance ratings were used in the main study, both sets of materials were pretested in two Navy recruiting districts, NRDs Detroit and St. Louis, with a total of 62 Navy recruiters.

In each location, the trial predictor battery was administered and reactions to the various inventories were elicited. Also administered during these sessions was a performance-rating scale package containing four sets of rating dimensions. The four sets of dimensions represent four different approaches to developing performance rating scales. Trait, behavior, multi-dimensional scaling, and factor analysis dimensions were included. The reasons for including so many scales in the pretest was to provide an opportunity to select strategies and individual scales that showed the most promise for yielding reliable and valid performance information. The scale development strategies and definitions for each dimension are presented in Appendix B.

Table 1

Measures Included in Pretest Predictor Battery

-
1. Personality Research Form (PRF) (Jackson, 1967)
Exhibition
Order
 2. California Psychological Inventory (CPI) (Gough, 1957)
Dominance
Sociability
Social Presence
Socialization
Achievement via Conformance
Good Impression
 3. Differential Personality Questionnaire (DPQ) (Tellegen, 1976)
Social Closeness
Hard Work
Authoritarianism
Impulsiveness
 4. Sales Effectiveness Scale (SES) (Dunnnette, 1976)
 5. Self-Description Inventory (SDI) (Ghiselli, 1954)
Intelligence
Supervisory Qualities
Initiative
Self-Assurance
Perceived Occupational Level
Decision-Making Ability
Sociometric Popularity
Perceived Maturity
 6. Strong-Campbell Interest Inventory (SCII) (Campbell, 1974)
 7. Biographical and Opinion Survey
-

In both NRDs three sets of ratings were gathered. Recruiters were asked to provide both self and peer evaluations. The peer ratings were directed toward all recruiters stationed in the same office. Supervisory evaluations were made by persons selected from the Zone Supervisor, Chief Recruiter, Enlisted Processing Officer, Executive Officer, and Commanding Officer positions. Those supervisory personnel who were available and seemed most qualified to rate recruiter performance provided the supervisory ratings. Two supervisor ratings were obtained for each recruiter whenever possible.

In addition to self, peer, and supervisory ratings, objective performance information was gathered in the form of the number of individuals each recruiter had enlisted in the Navy over the 6-month period from May

to October 1976. Such performance data were available for 54 of the 62 Detroit and St. Louis Navy recruiters. This information was prepared for analysis by first standardizing the production indices according to NRD. That is, standard scores were developed for each recruiter for each month by standardizing each month's production data within NRD. This operation corrects for unequal opportunity to bring persons into the Navy because of geographical location of the NRD, a practice that appears warranted based upon results of Arima's (1977) recent work.

Thus, for the Detroit and St. Louis pretest sample of 62 Navy recruiters, there were available for each recruiter performance ratings on 28 different rating scales, production data over a 6-month period ($N = 54$), and predictor information based on recruiter responses to the trial predictor battery.

Pretest Results

Performance dimension ratings were analyzed to assess interrater agreement along with leniency, restriction-of-range, and halo errors. Additionally, a factor analysis was conducted to summarize the performance rating information. The results of these analyses are presented in Table 2. Finally, to assess the stability of the production index chosen as an objective measure of recruiter performance, its reliability or consistency was examined across the 6-month period for which data were available. The following paragraphs describe the methods and objectives of these tasks.

Interrater Agreement

Intraclass correlations were computed depicting the interrater reliability of all possible combinations of self, peer, and supervisory ratings. Since the magnitude of the intraclass reliability coefficients depends to some extent on the number of raters per ratee, the coefficients obtained were adjusted to the two-rater level using the Spearman-Brown Prophecy Formula. This adjustment makes the reliabilities comparable and thus provides a means of assessing the relative degree of reliability associated with each combination of rater sources.

In general, the self-supervisor ratings, when pooled together, are most reliable, taking the number of raters per ratee into account. For deciding which ratings to use as criteria, however, the most appropriate index is the unadjusted intraclass coefficient. This is because the unadjusted intraclass coefficient reflects the reliability of the mean ratings, and these mean ratings are the data that are actually used as the criterion measures. Using this index of reliability, the peer and supervisory ratings and the ratings pooled across all three sources provide maximum reliability. Further, factor analysis of the self-peer-supervisor ratings pooled together, which is described later in this section, showed the clearest factor structure. Therefore, self, peer, and supervisory evaluations were employed in subsequent analyses. Reliabilities of the ratings from these three sources combined appear in Table 2.

Table 2

Pretest Rating Dimension: Results of Analyses

Rating Dimensions	Reliability ^a		Means ^b	Standard Deviations	Factor Loadings ^c		
	All Raters	Two-Rater Reliability			I	II	III
<u>Trait Scales</u>							
1. Initiative	.71	.55	6.37	1.41	.26	<u>.73</u> ^d	.36
2. Judgment	.56	.37	5.87	1.23	.46	<u>.65</u>	.26
3. Imagination and Originality	.69	.51	5.61	1.24	.57	.57	.35
4. Personal Impact	.51	.34	6.16	1.16	.59	.28	.56
5. Confidence	.63	.45	6.15	1.19	<u>.68</u>	.43	.30
6. Achievement Orientation	.70	.53	6.09	1.20	.32	.55	.56
7. Organization	.61	.43	5.50	1.22	.31	<u>.75</u>	.18
8. Wariness	.53	.35	6.45	1.13	.49	.00	<u>.67</u>
<u>Behavior Scales</u>							
9. Locating and Contacting Qualified Prospects	.63	.45	5.66	1.22	<u>.60</u>	.35	.41
10. Gaining and Maintaining Rapport	.44	.27	6.43	1.00	<u>.67</u>	.30	.51
11. Obtaining Information from Prospects and Making Good Person-Navy Fits	.66	.48	5.98	1.19	<u>.64</u>	.52	.33
12. Salesmanship Skills	.73	.56	6.07	1.42	<u>.80</u>	.43	.26
13. Establishing and Maintaining Good Relationships in the Community	.60	.42	5.88	1.19	.24	.33	<u>.73</u>
14. Providing Knowledgeable and Accurate Information about the Navy	.61	.43	6.34	1.15	.38	<u>.74</u>	.26
15. Administrative Skills	.62	.45	5.42	1.28	.39	<u>.71</u>	.23
16. Supporting Other Recruiters and the Command	.36	.21	5.65	1.09	.22	<u>.64</u>	.50

^a Interrater agreement indices for self, peer, and supervisory ratings.

^b Means and standard deviations of ratings pooled across the self, peer, and supervisory rating sources.

^c Varimax-rotated factor analysis solution of pretest performance ratings. Three factors--selling, planning, and community relations--emerged. The Overall Effectiveness dimension was not included, since one intent of the analysis was to summarize the data for each dimension.

^d The loadings in boxes indicate dimensions that were used in interpreting a factor.

Table 2 (Continued)

Rating Dimensions	Reliability ^a		Means ^b	Standard Deviations	Factor Loadings ^c		
	All Raters	Two-Rater Reliability			I	II	III
<u>NDS Scales</u>							
17. Prospecting	.57	.39	6.12	1.14	.61 ^d	.40	.28
18. Gathering Information about Applicants	.39	.23	5.95	.97	.23	.61	.38
19. Planning and Organizing Own Recruiting Schedule and Practices	.62	.44	5.37	1.24	.47	.70	.18
20. Expending Extra Effort to Aid Applicants or Recruits	.43	.27	6.41	.99	.21	.40	.68
21. Preparing Recruits and Their Parents for Navy Life	.33	.19	6.38	1.12	.36	.44	.62
22. Salesmanship	.77	.61	6.21	1.38	.75	.47	.29
23. Developing Productive Relationships in the Community	.53	.35	5.31	1.31	.23	.38	.65
24. Gaining and Maintaining Friendly Relationships with Prospects	.48	.30	6.50	1.07	.60	.13	.58
<u>Factor Scales</u>							
25. Prospecting Skills	.61	.43	5.79	1.09	.68	.49	.31
26. Selling Skills	.70	.53	6.23	1.28	.80	.45	.24
27. Human Relations Skills	.64	.46	5.90	1.34	.57	.47	.47
28. Overall Effectiveness as a Recruiter	.81	.67	5.83	1.47	--	--	--
Median Reliability	.61	.43					
Percent Common Variance Accounted for					27	26	20

^a Interrater agreement indices for self, peer, and supervisory ratings.^b Means and standard deviations of ratings pooled across the self, peer, and supervisory rating sources.^c Varimax-rotated factor analysis solution of pretest performance ratings. Three factors—selling, planning, and community relations—emerged. The Overall Effectiveness dimension was not included since the intent of the analysis was to summarize the data for each dimension.^d The loadings in boxes indicate dimensions that were used in interpreting a factor.

Leniency

As Borman, Dunnette, and Johnson (1974) and Schwab, Heneman, and DeCotiis (1975) among others have pointed out, one cannot directly assess leniency unless the "true" criterion scores are available against which to compare the level of ratings. For example, very high ratings may reflect actual high-level performance of ratees rather than inflated depictions of these individuals' performance effectiveness. Yet, it can generally be inferred that mean ratings near the scales' highest point indicate the presence of leniency. Table 2 contains the mean ratings for each of the scales. The means of 5.31 to 6.50 on a 9-point scale suggest that the ratings are probably reasonably free from leniency error.

Restriction of Range

The standard deviations in Table 2 provide some indication of the restriction-of-range error. Again, because the true range of performance on each scale is unknown, the exact degree of restriction of range cannot be determined; however, the standard deviations, which range from .97 to 1.47, suggest that the range of the ratings is at a reasonable level.

Halo

Assessing precisely the degree of halo error is impossible unless the true degree of correlation between dimensions is known. Then the observed correlation between ratings on the various dimensions can be compared to the true correlations, with higher observed correlations indicating halo error. Since there is no way to know the true correlations between dimensions, error cannot be precisely evaluated; however, the magnitude of the observed correlations (mean dimension intercorrelations = .63) suggests that some halo error is present in these ratings. Though actual performance on these dimensions is probably correlated to some extent, almost certainly it does not correlate as highly as these results suggest. This means that raters were forming general judgments about individual ratees and evaluating each ratee at much the same level on different dimensions.

Thus, overall, the leniency and restriction of range of these ratings are within acceptable limits and the reliability is reasonably high when ratings are pooled across the three sources—self, peer, and supervisors. However, a pronounced halo error appears to pervade the ratings.

Determining Dimensionality of the Ratings

Ratings from the peer-supervisor and self-peer-supervisor sources were pooled, and the ratings from all of the dimensions except Overall Effectiveness were intercorrelated and factor-analyzed. The three-factor, varimax-rotated solution from the self-peer-supervisor sources proved to be most readily interpretable. The factor loadings appear in Table 2. The factors were named as follows:

- I. Prospecting and Selling.
- II. Planning, Organizing, and Administrative Skills.
- III. Developing Good Navy-Community Relations and Expending Extra Effort to Aid Prospects and Recruits.

Developing an Objective Production Index

The consistency of the production index was assessed by computing an intraclass correlation coefficient that compares variability in the production index within recruiter with total variability in production indices across recruiters for the 6-month period. This analysis-of-variance procedure yielded an intraclass correlation of .88, suggesting that the production figures are stable over time. (The .88 intraclass correlation can be interpreted as the reliability of the mean production index scores averaged across 6 months). Of course, high reliability does not ensure the relevance of this index as a performance criterion, but it is a compelling summary measure of "bottom line" recruiter performance.

Pretest Performance Criteria

The production index, along with scores on the three factors discussed previously, provided four performance criteria against which to assess the validity of the predictor inventory scales and items. A fifth criterion--the overall performance rating dimension--was selected because of its high reliability ($r = .81$) and its conceptual appropriateness as a summary performance indicator.

Table 3 presents the intercorrelations between these five criteria and the reliability of each. The overall performance rating correlates highly with the objective production index, but neither of these summary performance measures correlates highly with effectiveness in individual facets of the job as represented by scores on the three factors. This result makes good sense. Two overall performance measures might be expected to correlate more highly with each other than with scores on individual aspects of performance. Further, Table 3 indicates that the reliabilities of the factor scores are reasonably high.

Overall, the five criteria appear to reflect recruiter performance effectively. The three uncorrelated composite dimensions represent three conceptually meaningful aspects of performance, and the objective production index and the overall performance rating provide two summary effectiveness measures.

Using Pretest Data to Refine Rating Scale Materials

To develop a final set of performance rating scales for the main study, the content of the various rating scales, the factor structure of the three-factor varimax rotated solution discussed previously, and the reliabilities obtained for the ratings on each scale were examined. First, because of the considerable effort extended previously to develop behavior summary scales for Navy recruiters (Borman, Hough, & Dunnette, 1976), the eight behavioral performance scales presented in Table 2 (rating dimensions 9 to 16) were selected for use. These scales appeared to have good potential for providing performance ratings of a high quality. The rest of the scales (1 to 8 and 17 to 27) were considered as possible supplementary performance dimensions.

Table 3
Intercorrelations and Reliabilities of the Final
Five Performance Criteria--Pretest Data

Reliabilities	Performance Criteria	Intercorrelations				
		1	2	3	4	5
.69 ^a	1. Prospecting and Selling	--				
.65	2. Planning, Organizing, and Administrative Skills	.06	--			
.59	3. Developing Good Navy- Community Relations and Expending Extra Effort to Aid Prospects or Recruits	.04	.05	--		
.81	4. Overall Effectiveness as a Recruiter	.77	.46	.30	--	
.88	5. Production	.50	.19	.10	.70	--

^aReliabilities for the first three criteria index the stability of the factor scores associated with these dimensions.

The following criteria were used to select scales from the supplementary group of 19: (1) adequate interrater reliability; (2) high loading on one of the three factors; and (3) contribution to the conceptual completeness of the total list of dimensions. Using these guidelines for including extra scales, the following dimensions were selected: Initiative, Judgment, Confidence, Achievement Orientation, Organization, Warmth, Expending Extra Effort to Aid Applicants or Recruits, and Developing Productive Relationships in the Community.

The content of these scales appeared to supplement the content of the behavior summary scales, and each extra scale loaded relatively highly on one of the three factors from the pretest analysis, thereby raising the probability of a meaningful factor analysis solution in the main study. Thus, eight behavior summary scales and eight supplementary scales selected from pretest data were selected for use in the main part of the study. Additionally, Overall Effectiveness as a Recruiter was included to provide an overall performance scale. See Appendix C for the rating scale package used with the main Navy and Marine Corps samples.

Using Pretest Responses to Revise the Trial Predictor Battery

Pretest responses to the predictor inventories were used to refine the trial predictor battery. First, correlations between the personality scales from the battery and five pretest performance criteria were examined. These correlations appear in Table 4. Because many of Ghiselli's Self-Description Inventory (SDI) scales correlated near zero with the performance

Table 4
Correlations Between Pretest Predictor Battery
Personality Scales and the Five Pretest Performance Criteria

Performance Criteria	Prospecting and Selling	Planning and Organizing	Good Relations	Overall Effectiveness	Production
<u>PRF</u>					
1. Exhibition	.13	.11	.32	.22	.10
2. Order	-.08	.28	.08	.12	.16
<u>CPI</u>					
3. Dominance	.16	.23	.24	.32	.15
4. Sociability	.13	.16	.41	.24	.10
5. Social Presence	.10	-.01	.32	.14	.12
6. Socialization	-.18	.19	.36	.06	-.01
7. Achievement via Conformance	-.01	.21	.29	.17	-.05
8. Good Impression	.19	.25	.31	.32	.04
<u>DPQ</u>					
9. Social Closeness	.06	.14	.35	.18	.05
10. Hard Work	.06	.34	.30	.32	.23
11. Authoritarianism	-.06	.30	-.01	.06	.02
12. Impulsiveness	.13	-.28	-.20	-.13	-.15
<u>SES</u>					
13. Sales Effectiveness	.18	.05	.28	.29	.37
<u>SDI</u>					
14. Intelligence	-.06	-.19	.18	.13	-.05
15. Supervisory Qualities	.01	-.26	-.07	-.11	-.04
16. Initiative	-.08	.10	-.13	-.07	-.12
17. Self-Assurance	.05	-.18	.04	.01	.23
18. Perceived Occupational Level	-.18	-.04	-.05	-.16	-.21
19. Decision Making Ability	.09	.08	.22	.15	.03
20. Sociometric Popularity	.04	-.01	.09	.10	.11
21. Perceived Maturity	-.10	-.02	-.16	-.04	-.12

criteria, this instrument was eliminated from the battery with the exception of a single scale, Self-Assurance, which correlated .23 with the production index. To ease administration of this scale, only the "most descriptive adjective" items were used. In fact, that half of the scale correlated more highly with the objective production index than did the total scale ($r = .30$ vs. $r = .23$), suggesting that no validity would be lost by this change. Also, Dunnette's Sales Effectiveness Scale, which has the same format as the SDI, contains primarily most descriptive adjective pairs; therefore, those 32 Sales Effectiveness Scale items with the least descriptive format were eliminated. Like the elimination of the SDI items, this change provided no appreciable change in pretest validity. Altogether, the SDI and Sales Effectiveness section of the battery was reduced from 106 to 64 items based upon pretest information.

Since item-level analyses were to be performed on the Strong-Campbell Interest Inventory, it remained intact for the main battery administration effort.

Additionally, item-level analyses were to be performed on the Biographical and Opinion Survey; consequently, no items were eliminated at this time, though a few of the response alternatives were changed where it seemed warranted. This instrument and a sample revision appear in Appendix D.

Forming Predictor-Criterion Correlation Hypotheses Using Pretest Data

Data from the pretest, besides helping to refine the predictor battery and performance criterion materials, aided in formulating hypotheses about predictor-criterion relationships. These hypotheses are useful in estimating the cross-validity of the battery. Essentially, forming judgments about predictor-criterion relationships helps to eliminate items or scales from consideration as predictors for a criterion when their conceptual relationships with that criterion are weak or nonexistent. For example, Borman (1973) and Dunnette (1976a) have matched personality, ability, and interest scales to individual performance criteria according to the conceptual ties between predictor constructs and the performance dimensions. These studies employed the Monte Carlo procedure described in Appendix E to assess the empirical validity of predictor composites formed by combining the scales selected.

The approach used in the present study extends this strategy by using both conceptual ties and empirical data to help make judgments about these relationships. It estimates the cross-validity of various parts of the battery in one of two ways: (1) by applying the preselected keys directly in the main administration samples; or (2) by performing the Monte Carlo analysis. Reasons for these choices appear in the next section.

Accordingly, pretest correlations between predictors and criteria were examined closely and the conceptual links between each predictor scale or item and each performance criterion¹ were studied to arrive at the list of

¹The criteria referred to here are pretest criteria, not main sample criteria. Thus, these hypotheses were formed for the three performance criteria generated by the pretest factor analysis. However, the criterion analysis of the main sample data indicated substantially the same underlying factors, thereby making the hypothesized criterion relationships appropriate for main sample data.

hypothesized predictor-criterion linkages presented in Appendix F. Thus, both a simple cross-validation model and the Monte Carlo procedure developed by Rosse and used by Borman (1973), Dunnette and Motowidlo (1975), Rosse and Hellervik (1975), and Dunnette (1976b) were selected for use with the main sample in this study.

Assigning Cross-Validation Strategies to Different Sections of the Predictor Battery

These predictor items and scales selected on the basis of pretest information were separated (for each criterion) into three predictor sets--personality, interest, and biographical. The next step was to decide where Monte Carlo item analysis procedures should be followed and where simply cross-validating the preselected keys on the main sample was more appropriate.

Personality Scales

The hypothesized personality predictors, containing variables at the scale level, did not appear to be good candidates for Monte Carlo item analysis treatment. The pretest results indicated that several personality scales correlated significantly with one or more criteria and that the pattern of relationships between scales and criteria was conceptually reasonable. Thus, it was decided to simply cross-validate the personality composites as developed from pretest data with the criteria toward which they were targeted.

SCII

The Monte Carlo procedure appeared appropriate for investigating relationships between the SCII and each of the five criteria for the Navy sample. This is because relationships between SCII scales (i.e., basic, occupational, and Holland theme scales) and the criteria in the pretest data were not very promising; thus, some kind of item analysis seemed warranted. And also, the main sample size for the Navy ($N = 267$) was large enough that the Monte Carlo item analysis procedure should create relatively stable keys for each run, thereby ensuring reasonably accurate estimates of the cross-validity of the item keys. See Appendix E for a further description of the Monte Carlo procedure.

For the Marine Corps sample ($N = 118$), the sample size appeared to be too small to provide stable keys for each Monte Carlo run. Therefore, it seemed preferable to use a simple cross-validation model to estimate the validity of keys selected on the basis of pretest data. Accordingly, for each criterion, SCII items were selected and keyed (i.e., a direction was specified) that showed promise, both conceptually and empirically (pretest data), for predicting performance. These keys were then cross-validated on the main sample ($N = 118$).

Biographical Survey

The relatively small number of biographical survey items selected for cross-validation suggested that the Monte Carlo procedure might be

appropriate for both samples. Even for the Marine Corps group, the small number of items included in the cross-validation item pools meant that relatively few invalid items would enter the keys on individual Monte Carlo runs. Therefore, the Monte Carlo procedure was expected to provide a reasonably accurate estimate of the keys' cross-validities.

Description of the Sample for the Main Study

For the main administration of the inventory and performance rating materials, a geographically representative sample of 267 Navy recruiters from ten Navy Recruiting Districts (NRDs) was selected. A list of those NRDs and the number of Navy recruiters participating from each is shown in Figure 1, which presents the same information for the Marine Corps sample. While offering a good spread geographically, neither the Navy nor the Marine Corps sample is representative in terms of a rural, suburban, inner city split. The inner city and suburban recruiter population is overrepresented, though recruiters canvassing in rural areas were included in the Amarillo, El Paso, and Oklahoma City groups.

A sampling plan which underrepresents persons recruiting in rural areas was used here for two reasons. First, countrywide, inner city, and suburban recruiters as a group comprise the vast majority of the recruiting force, and they recruit many more individuals into the Navy and Marine Corps than do their counterparts working rural areas. Second, the cost of collecting data would have been increased considerably with any substantial inclusion of recruiters from rural stations.

Overall, the samples appear to be representative of the Navy and Marine Corps suburban and inner city recruiter populations.

Procedures for the Main Study

To collect predictor inventory and performance ratings, meetings were arranged with small groups of no more than 16 recruiters at a time. This procedure provided maximum personal attention to participating recruiters during the performance-rating part of the program. The small group sessions were designed to facilitate the training of raters and to provide considerable opportunity for questions about the rating form, the inventory battery, and the study in general.

Each session began with a 10- to 15-minute briefing that outlined the purpose of the study and explained the procedures to be followed during the 3-hour session. After the briefing, the group was divided into two roughly equal-sized subgroups, and one subgroup began the performance ratings while the other subgroup started completing the inventories. After the first subgroup completed the performance ratings, the recruiters began completing the trial inventories. The other subgroup performed the rating task after completing the inventories.

In the performance rating session, participants were asked to read the directions for the rating scales (see Appendix C for a copy of the directions and of the rating scales) and provided an example of how to rate an individual on the first dimension, Locating and Contacting Qualified Prospects. They were allowed to ask questions and were instructed to

	<u>Navy Recruiters</u>	<u>Marine Corps Recruiters</u>
<u>Area 1</u>		
Philadelphia	28	15
<u>Area 3</u>		
Southeast region (includes Jacksonville, Miami, Tampa)	33	17
<u>Area 4</u>		
Washington, D. C.	33	20
<u>Area 5</u>		
Minneapolis	--	12
Chicago	33	20
<u>Area 7</u>		
Houston	26	--
Southwest (includes Amarillo, Albuquerque, and El Paso)	25	8
Oklahoma City	22	7
<u>Area 8</u>		
Oakland	35	10
Los Angeles	32	9
Total	267	118

Figure 1. Description of main samples according to Navy Area Recruiting Districts.

evaluate themselves and their station mates who were participating in the research program.

In a separate rating session, performance evaluations were gathered from supervisors selected from the supervisory levels--Chief Recruiter, Enlisted Processing Officer, and Zone Supervisor (Noncommissioned Officer in Charge or, in the Marine Corps, Sergeant Major). As with the pretest, two supervisory ratings were gathered for each recruiter whenever possible.

In the testing sessions, recruiters were first briefed about the content of the inventories to be completed, then asked to begin responding to the trial inventory battery at their own speed.

Finally, production statistics were gathered for recruiter participants in the form of the number of persons each had brought into the Navy or Marine Corps from May through October of 1976. Thus, the main data collection effort yielded self, peer, and supervisory ratings; production information; and responses to the revised inventory battery.

RESULTS

Criterion Development Results

The performance ratings gathered from supervisors and recruiters were, as in the pretest, analyzed to evaluate various rating errors. Interrater reliability of the ratings was assessed along with leniency, restriction-of-range, and halo errors. Then, the dimensionality of the performance ratings was examined in order to create a final set of criteria intended to comprehensively and efficiently reflect recruiter effectiveness.

Interrater Agreement

Table 5 shows the intraclass correlation coefficients indicating the interrater reliability of ratings from the self, peer, and supervisory sources pooled together. Results are displayed for the Navy and Marine Corps samples. As in the pretest, the reliabilities of ratings from all combinations of these sources were examined and the peer-supervisor and self-peer-supervisor groupings were found to provide the most reliable mean ratings. The self-peer-supervisor grouping was selected because the factor structure of ratings from this combination of sources was more readily interpretable than the factor structure obtained from the peer-supervisor ratings. In general, reliability is acceptable for the Navy (median $r = .57$) but marginal for the Marine Corps (median $r = .48$) scales.

Leniency

As was explained earlier, the rating errors of leniency, restriction of range, and halo can be evaluated only indirectly. As for leniency, the mean ratings on individual dimensions in the Navy sample range from 6.86 to 7.81 on the 10-point behavior dimensions and from 6.23 to 7.01 on the 9-point supplementary scales. For the Marine Corps sample, the ranges are 6.55 to 7.83 on the 10-point and 5.79 to 6.92 on the 9-point scales. See Table 6 for a complete presentation of the means. These mean ratings are far enough from the scales' highest points that leniency does not appear to be a serious problem.

Restriction of Range

The standard deviations of the ratings for each dimension are also depicted in Table 6. They indicate that a reasonable degree of variation in effectiveness is present in the rated performance of the recruiters in the two samples.

Halo

The magnitude of the correlations between dimensions is in general high. In the combined Navy and Marine Corps sample, for example, these correlations range from .35 to .81 for the ratings derived from the three sources pooled together. The median interdimension correlation is .62. The magnitude of these relationships suggests that some degree of halo error is present in the ratings.

Table 5
Interrater Agreement Indices for Self, Peer, and Supervisor Ratings

Rating Dimensions	Reliabilities			
	Navy		Marine Corps	
	All Raters	Two-Raters	All Raters	Two-Raters
1. Locating and Contacting Qualified Prospects	.73	.57	.67	.54
2. Gaining and Maintaining Rapport	.54	.37	.48	.35
3. Obtaining Information from Prospects and Making Good Person-Navy Fits	.46	.30	.40	.28
4. Salesmanship Skills	.66	.49	.58	.45
5. Establishing and Maintaining Good Relationships in the Community	.57	.40	.44	.32
6. Providing Knowledgeable and Accurate Information about the Navy	.47	.30	.45	.32
7. Administrative Skills	.57	.40	.45	.33
8. Supporting Other Recruiters and the Command	.53	.36	.42	.30
9. Initiative	.65	.48	.54	.41
10. Judgment	.64	.47	.34	.23
11. Confidence	.67	.50	.57	.43
12. Achievement Orientation	.59	.42	.61	.48
13. Organization	.61	.43	.50	.37
14. Warmth	.52	.35	.39	.27
15. Expending Extra Effort to Aid Applicants or Recruits	.47	.30	.47	.35
16. Developing Productive Relationships in the Community	.53	.36	.48	.35
17. Overall Performance	.78	.64	.66	.53
Median Reliability	.57	.40	.48	.35

Table 6

Means and Standard Deviations of Ratings Pooled Across the
Self-Peer-Supervisor Rating Sources

Rating Dimensions	Navy		Marine Corps	
	Mean	Standard Deviation	Mean	Standard Deviation
1. Locating and Contacting Qualified Prospects	6.86	1.26	6.66	1.30
2. Gaining and Maintaining Rapport	7.76	1.07	7.64	1.15
3. Obtaining Information from Prospects and Making Good Person-Navy Fits	7.25	.97	7.13	1.14
4. Salesmanship Skills	7.26	1.25	7.27	1.30
5. Establishing and Maintaining Good Relationships in the Community	7.21	1.17	6.58	1.25
6. Providing Knowledgeable and Accurate Information about the Navy	7.81	1.04	7.83	1.12
7. Administrative Skills	7.08	1.21	6.55	1.34
8. Supporting Other Recruiters and the Command	7.55	1.13	7.15	1.25
9. Initiative	6.66	1.17	6.48	1.14
10. Judgment	6.43	1.08	6.50	.90
11. Confidence	6.82	1.15	6.92	1.05
12. Achievement Orientation	6.67	1.06	6.53	1.12
13. Organization	6.23	1.06	5.79	1.04
14. Warmth	6.93	1.04	6.90	1.01
15. Expending Extra Effort to Aid Applicants or Recruits	7.01	.96	6.86	.99
16. Developing Productive Relationships in the Community	6.34	1.08	6.07	1.14
17. Overall Performance	6.50	1.31	6.40	1.24

Determining Dimensionality of the Ratings

Although halo did appear to pervade the ratings, an effort was made to identify factors that might underlie the relationship between dimensions. The Navy and Marine Corps ratings were pooled and a principal-components factor analysis was applied to the intercorrelations among dimensions for each of the following rating source groupings: (1) peer-supervisor; and (2) self-peer-supervisor. The eigenvalues of the two solutions were examined and the two-, three-, and four-factor, varimax-rotated solutions were studied for their psychological meaningfulness. The three-factor solution for the self-peer-supervisor ratings appeared most meaningful. That solution appears in Table 7. The following factors were named:

I. Selling Skills: selling the Navy or Marine Corps effectively to prospects; displaying confidence and effectiveness in the recruiting sequence--prospecting, selling, and closing.

II. Human Relations Skills: establishing and maintaining good interpersonal relations with prospects and recruits, persons in the community, fellow recruiters, and supervisory personnel.

III. Administrative Skills: planning ahead and organizing time efficiently; completing paperwork accurately and on time.

The content of these factors agrees closely with the content of the pretest factors, suggesting that the three-factor solution represents a stable picture of the underlying dimensions of performance on the Navy or Marine Corps recruiter job. Further, the main Navy and the Marine Corps sample ratings were analyzed separately to evaluate the comparability in factor structure across two groups. The two three-factor solutions were markedly similar when each was interpreted by itself. In addition, a crude numerical index of the comparability in these factor solutions was provided by correlating the factor loadings for the three pairs of varimax-rotated factors. Those correlations were .83, .69, and .86, suggesting further the stability of these orthogonal dimensions. Finally, the reliabilities of the factor scores are reasonably high. Table 8 indicates that these reliabilities range from .51 to .68 for the two samples.

In summary, based on conceptual considerations, reliability results, and results of the factor analyses, all ratings--self, peer, and supervisor ratings--were selected for subsequent steps in the study.

Developing the Objective Production Index

As in the pretest study, monthly production data were available for many recruiters in the sample for the 6 months from May through October of 1976 (i.e., number of accessions for each of the 6 months). The stability of the corrected production index, adjusted within district in the same way as was pretest production, was evaluated by computing its reliability using methods identical to those employed with pretest data. The intraclass correlation coefficients representing the reliability of the mean corrected production scores across the 6 months were .73 for the Navy sample and .83 for the Marine Corps sample. These

Table 7

Three-Factor Varimax Rotated Solution for Self-Peer-Supervisor
Ratings Pooled Together: Navy and Marine Corps Main Sample Combined

Rating Dimensions ^a	I. Selling Skills	II. Human Relations Skills	III. Administrative Skills
1. Locating and Contacting Qualified Prospects	<u>.62</u> ^b	.44	.40
2. Gaining and Maintaining Rapport	.57	.60	.21
3. Obtaining Information from Prospects and Making Good Person-Navy or Marine Corps Fits	<u>.55</u>	.39	.43
4. Salesmanship Skills	<u>.77</u>	.35	.32
5. Establishing and Maintaining Good Relationships in the Community	.26	<u>.68</u>	.31
6. Providing Knowledgeable and Accurate Information about the Navy or Marine Corps	.49	.41	.44
7. Administrative Skills	.27	.25	<u>.75</u>
8. Supporting Other Recruiters and the Command	.39	.52	.39
9. Initiative	.59	.39	.56
10. Judgment	.57	.31	.56
11. Confidence	.64	.41	.39
12. Achievement Orientation	.53	.49	.48
13. Organization	.31	.32	<u>.78</u>
14. Warmth	.40	<u>.63</u>	.14
15. Expending Extra Efforts to Aid Applicants or Recruits	.32	<u>.71</u>	.32
16. Developing Productive Relationships in the Community	.25	<u>.74</u>	.34
Percent Common Variance Accounted For	25	25	21

^aThe Overall Performance dimension was not included because the intent of the analysis was to examine the dimensionality of scales associated with individual aspects of the job.

^bThe loadings in boxes indicate dimensions that were used in interpreting a factor.

Table 8

Intercorrelations and Reliabilities of the Five Criteria

Reliabilities	Criteria	Intercorrelations				
		1	2	3	4	5
Navy						
.68 ^a	1. Selling Skills	--				
.53	2. Human Relations Skills	.23	--			
.62	3. Administrative Skills	.21	.12	--		
.78	4. Overall Performance	.79	.50	.51	--	
.73	5. Production	.43	.15	.17	.52	--
Marine Corps						
.63	1. Selling Skills	--				
.55	2. Human Relations Skills	.03	--			
.51	3. Administrative Skills	.11	-.02	--		
.66	4. Overall Performance	.72	.42	.40	--	
.83	5. Production	.45	.31	.18	.59	--

^aReliabilities for the Selling, Human Relations, and Administrative Skills dimensions index the stability of the factor scores associated with these dimensions.

relatively high reliabilities indicate that the corrected accessions-per-month measure is at least consistent over time; recruiters who have high production one month tend to have high production during other months, and recruiters less productive in a given month tend to be relatively unproductive during other months as well. As discussed earlier, high reliability implies nothing about this measure's relevance as an index of job effectiveness, but the index does possess a certain face validity as an indicator of a recruiter's "bottom line" contribution to the Navy or Marine Corps recruiting effort.

Performance Criteria

The production index, along with the scores on these three factors identified by analyses, provided four performance criteria. As in the pretest, a fifth criterion--Overall Performance--was added to serve as a summary criterion.

Relationships Between Criteria

Table 8 depicts the correlations between the five criteria for the Navy and Marine Corps separately. As expected, the first three criteria are relatively uncorrelated because of the manner in which the factor scores were computed. As such, they represent effectiveness criteria in three distinctly different aspects of the Navy and Marine Corps recruiter job. The Overall Performance rating correlates moderately with the Human Relations Skills and Administrative Skills criteria, and highly ($r_s = .79$ and $.72$) with Selling Skills. This pattern of correlations appears reasonable because the Selling Skills criterion defines the heart of the recruiting sequence--prospecting, selling, and closing--and therefore should be the criterion most closely related to overall performance. The same observation holds for the Production criterion. Of the three performance criteria associated with individual facets of the job, it relates most highly to Selling Skills. Further, the correlations between this measure and the Overall Performance rating is at least moderate ($r_s = .52$ and $.59$) and is higher than the relationships between production and performance in any single facet of the job.

Thus, the five criteria possess adequate to high reliability, the criteria measuring performance in individual facets of the job appear to reflect well three relatively independent aspects of Navy or Marine Corps recruiter job performance, and the pattern of correlations between the facet criteria and the two measures of overall effectiveness make good conceptual sense. These criterion development results pave the way for attempts to predict performance on each criterion measure using information on differences between recruiters that is available from predictor data.

Trial Predictor Battery Validation Results

This section describes relationships between the five performance criteria and predictor composites developed according to the hypotheses about predictor-criterion linkages. The composites can be found in Appendix F. Additionally, it presents data that can be used to evaluate racial bias and the differential validity of these predictor composites.

First, Tables 9 and 10 show the correlations between personality scales and each of the five criteria for the Navy ($N = 267$) and the Marine Corps ($N = 118$) separately. For the Navy sample, the relationships are not as high in general as was found when pretest data were analyzed. Marine Corps results are similar (both in magnitude and pattern of the correlations) to the results obtained from the pretest data.

Cross-Validation Results for Preselected Keys

Results of the personality composite cross-validation are shown in Table 11. Administrative Skills is clearly the criterion most readily predicted by the personality composites for both the Navy and the Marine Corps samples. The Production criterion, though reasonably well predicted in the Marine Corps sample, correlated near zero with the Production composite within the Navy group.

Table 12 displays results of correlations between each of the five criteria and the SCII keys for the Marine Corps groups. The keys were formed on the basis of pretest information. These cross-validity coefficients hover about the .20 to .25 level with the exception of the relationship between Production and the Production SCII predictor key, for which $r = .10$.

Monte Carlo Cross-Validation Results

For the Monte Carlo analyses, only those items preselected on the basis of pretest data and conceptual considerations were included in the pool that entered Monte Carlo runs. Further, the direction (i.e., positive or negative) of many of the relationships was specified ahead of time. Identifying the intended direction prevents an item from entering a Monte Carlo run when it meets the item inclusion criterion in the wrong direction. The direction and criterion limit options of the Monte Carlo computer program are explained more fully in Appendix E.

Results of the Monte Carlo analyses for the SCII (Navy sample) appear in Table 13. The "total sample" validity coefficients represent the correlation between each criterion and the keys formed from total sample data ($N = 267$). The cross-validity coefficients are median Monte Carlo estimates of these keys' validities. They are probably underestimates of the validities because the keys developed for each Monte Carlo run are formed on only half the total sample and are therefore likely to be less stable than the final keys developed using the entire sample.

Table 14 depicts Monte Carlo results for the Biographical Survey data. In general, the results are disappointing. Only with the criteria Administrative Skills and Overall Performance in the Navy sample do cross-validities reach even the .20s. For the Marine Corps groups, the validities are only .15. Again, these Monte Carlo validities are probably underestimates of the final keys' true validities.

Table 9

Correlations Between Predictor Battery Personality Scales
and the Five Performance Criteria: Main Sample, Navy

Criteria	Selling Skills	Human Relations Skills	Administrative Skills	Overall Performance	Production
1. Exhibition	.20	.06	-.02	.16	-.03
2. Order	-.08	-.08	.30	-.02	-.04
3. Dominance	.18	.04	.17	.15	-.05
4. Sociability	.08	.06	.01	.07	-.08
5. Social Presence	.20	.07	-.10	.10	-.03
6. Socialization	.01	.07	.27	.11	.12
7. Achievement via Conformance	.01	.03	.24	.06	.02
8. Good Impression	-.02	.11	.16	.06	-.02
9. Social Closeness	.07	.20	.05	.17	.04
10. Hard Work	.06	.12	.14	.15	.03
11. Authoritarianism	-.11	-.06	.01	-.03	-.06
12. Impulsiveness	.14	.07	-.20	.07	.05
13. Sales Effectiveness	.17	.09	.08	.21	.02
14. Self-Assurance	.00	.00	.19	.09	-.02

Table 10

Correlations Between Predictor Battery Personality Scales
and the Five Performance Criteria: Main Sample, Marine Corps

Criteria	Selling Skills	Human Relations Skills	Administrative Skills	Overall Performance	Production
1. Exhibition	.10	.04	-.17	.03	-.25
2. Order	-.01	.03	.40	.12	.14
3. Dominance	.10	.22	.07	.13	.00
4. Sociability	.06	.14	-.05	.09	.00
5. Social Presence	.16	.04	-.11	.09	-.08
6. Socialization	.06	.11	.35	.20	.17
7. Achievement via Conformance	.12	.14	.32	.17	.14
8. Good Impression	.04	.18	.34	.20	.09
9. Social Closeness	.06	.17	.01	.09	-.09
10. Hard Work	.08	.19	.24	.22	.23
11. Authoritarianism	.03	.22	.24	.23	.15
12. Impulsiveness	-.22	-.01	-.38	-.23	-.25
13. Sales Effectiveness	.27	.08	.13	.31	.30
14. Self-Assurance	.05	-.03	.06	.04	-.08

Table 11
Cross-Validation Results for the Personality Composites

Criteria	Navy	Marine Corps
Selling Skills	.21**	.19*
Human Relations Skills	.16**	.19*
Administrative Skills	.29**	.38**
Overall Performance	.21**	.24**
Production	.01	.24*

*p < .05

**p < .01

Table 12
Cross-Validation Results for the SCII Keys
(Marine Corps Only)^a

Criteria	Marine Corps
Selling Skills	.19*
Human Relations Skills	.24**
Administrative Skills	.20*
Overall Performance	.26**
Production	.10

^aSCII results for the Navy were obtained using the Monte Carlo cross-validation procedure. Those results are contained in Table 13.

*p < .05

**p < .01

Table 13

Monte Carlo Results—SCII, Navy Sample

Criteria	Total Sample Validity Coefficients	Cross-Validity Coefficients
Selling Skills	.32**	.19**
Human Relations Skills	.23**	.17**
Administrative Skills	.38**	.26**
Overall Performance	.33**	.17**
Production	.11	-.02

**p < .01

Table 14

Monte Carlo Results—Biographical Survey

Criteria	Navy		Marine Corps	
	Total Sample Validity Coefficients	Cross- Validity Coefficients	Total Sample Validity Coefficients	Cross- Validity Coefficients
Selling Skills	.35**	.16**	.40**	-.01
Human Relations Skills	.30**	.05	.46**	.10
Administrative Skills	.42**	.23**	.49**	.15
Overall Performance	.39**	.23**	.53**	.15
Production	.30**	.01	.46**	.08

**p < .01

Estimated Validity for the Final Composites Pooled Together

To estimate the validity of the various predictor composites taken together, a simple procedure was employed that uses the cross-validity coefficients and intercorrelations between the predictor composites. The following formula was applied:

$$V_t = \frac{\sum_{i=1}^m w_i r_{iy}}{\sqrt{\sum_i \sum_j w_i w_j r_{ij}}}$$

where V_t = estimated cross-validity of the combined composites,

r_{iy} = estimated cross-validity of the i th predictor composite with the criterion y ($i = 1, 2, \dots, m$),

w_i and w_j = the weights applied, and

r_{ij} = the intercorrelation of the i th and j th predictor composites.

As weights, the standard deviations of composite scores were used. This procedure has the effect of giving extra weight to longer, and presumably more reliable, composites. Also, composites were eliminated from further consideration in this analysis when their validity did not appear sufficient to warrant inclusion in the final set of predictor measures (see Table 15 for the composites included).

Thus, the inventory responses for persons in the two samples, Navy (main sample) and Marine Corps, were scored according to the final keys developed for personality, SCII, and biographical composites. Then the composite scores were intercorrelated and these intercorrelations, along with the cross-validities, were employed to estimate the validity of the pooled predictor measures.

Results of this analysis appear in Table 15. Except for the Human Relations Skills and Production criteria in the Navy sample, estimated overall validities are uniformly in the .20s and .30s.

Black-White Comparisons: An Examination of Bias and Differential Validity in the Navy Sample

The Navy sample contained a large enough group of Black recruiters ($N = 36$) that an assessment of various black-white racial differences on the predictors, criteria, and predictor-criterion relationships was deemed technically feasible. (An N of 30 has been offered as a minimum sample size for validation research--Federal Register, 1976). The Marine Corps sample contained only 21 black recruiters, rendering any such comparisons technically inappropriate. Neither the Navy nor the Marine Corps samples contained sufficiently large numbers of women, Mexican-Americans, or any other minority group to make other kinds of intergroup comparisons possible.

Table 15

Intercorrelations of Predictor Composites/Keys and
Estimated Validity for the Final Composites Pooled Together

Predictor	Selling Skills			Human Relations Skills			Administrative Skills			Overall Performance			Production
	1	2	3	1	2	3	1	2	3	1	2	3	
Navy ^a													
1. Personality	--	--	--	--	--	--	--	--	--	--	--	--	N/A
2. SCII	.50	--	--	.69	--	--	.39	--	--	.63	--	--	
3. Biographical	.26	.09	--	--	--	--	.41	.20	--	.37	.24	--	
Validity of Composites Pooled Together	.24			.17			.31			.22			
Marine Corps ^b													
1. Personality	--	--	--	--	--	--	--	--	--	--	--	--	
2. SCII	.48	--	--	.65	--	--	--	--	--	.50	--	--	
Validity of Composites Pooled Together	.22			.22			.38			.27			.24

^a $r \geq .12$, $p < .05$; $r \geq .16$; $p < .01$.

^b $r \geq .18$, $p < .05$; $r \geq .24$, $p < .01$.

Table 16 shows the means and standard deviations on predictor composites and on the criteria for white ($N = 197$) and black ($N = 36$) Navy recruiters separately. These results suggest that mean predictor scores for the black and white recruiters in the sample are very similar. The largest differences occur for the Selling Skills biographical key, on which black Navy recruiters score approximately one-third of a standard deviation higher than their white counterparts. Differences in means or standard deviations are very slight in other cases. Except for Administrative Skills, black recruiters are rated as a little more effective than white recruiters, but again, the differences are not great--one-fourth standard deviation at the most.

Table 16

Black-White Differences: Means and Standard Deviations
on Predictor Composites and Performance Criteria

Variables	Criteria							
	Selling Skills		Human Relations Skills		Administrative Skills		Overall Performance	
	M	SD	M	SD	M	SD	M	SD
<u>Personality</u>								
Black	205.26	26.06	365.66	59.44	312.94	51.59	308.71	34.97
White	202.26	26.38	361.95	63.46	305.00	50.84	305.11	39.34
<u>SCII</u>								
Black	22.33	16.69	20.78	20.66	5.61	8.91	21.22	17.67
White	21.78	15.36	13.42	20.74	7.87	8.75	19.35	15.48
<u>Biographical</u>								
Black	-.17	1.11	-- ^a	-- ^a	-.23	2.22	-.66	2.11
White	-.58	1.32	-- ^a	-- ^a	-.61	1.78	-1.09	1.98
<u>Actual Performance</u>								
Black	51.04 ^b	--	52.01	--	50.96	--	6.76 ^b	--
White	48.38	--	49.86	--	51.74	--	6.39	--

^aBiographical items did not yield a scale for Human Relations Skills.

^bThe mean performance scores are T-scores (mean = 50, SD = 10) for all criteria except Overall Performance. Mean scores for that criterion are mean ratings on the Overall Performance dimension.

Table 17 shows the validity coefficients for the black and white recruiter groups separately. As with the means and standard deviations, few substantial differences can be noted between the two samples even though the small size of the black group makes the validity coefficients for the

black sample quite unstable. Only Selling Skills shows consistent differences, with the validity higher in the white recruiter group. However, even these differences are not statistically significant and are not particularly striking.

Table 17

Black-White Differences: Validity Coefficients
for Final Predictor Composites

Predictor	Selling Skills		Human Relations Skills		Administrative Skills		Overall Performance	
	Blacks	Whites	Blacks	Whites	Blacks	Whites	Blacks	Whites
Personality	.07	.21	.28	.16	.44	.16	.32	.20
SCII ^a	.16	.35	.19	.25	.33	.38	.32	.34
Biographical ^a	.07	.37	—	--	.59	.38	.40	.39

^aValidity coefficients for the SCII and Biographical Survey are total sample correlations, not cross-validities.

Overall, these data suggest that the final predictor composites will not show bias against either white or black candidates for the Navy recruiter job.

CONCLUSIONS

Performance Measurement: Differences Between the Present Study and Previous Efforts

During this project, considerable effort went into the development of criteria for measuring Navy and Marine Corps recruiter effectiveness. Briefly, the investigation of performance criteria included studying (1) various combinations of rating sources (i.e., self, peer, and supervisors), (2) the dimensionality of military recruiter effectiveness, and (3) the relationship between recruiter production (i.e., number of persons brought into the Navy or Marine Corps) and performance in the various aspects of the job. These steps were taken in order to arrive at final performance criteria that reflected relevant, reliable, comprehensive, and yet parsimonious measurement of recruiter performance. The relatively elaborate criterion development procedures may well have led to the reasonably good validities obtained in the study.

Reviewing briefly the criterion development approaches taken in other studies of the military recruiter job should highlight the differences between those approaches and the one pursued here. The three types of measures used in these previous studies are (1) uncorrected production; (2) adjusted or corrected production; (3) ratings, ranking, or nominations. Yet, each of these methods may fail to describe accurately a military recruiter's true effectiveness.

First, production, though it seems so legitimate as a "bottom line" index of a recruiter's effectiveness, is probably riddled with the kinds of errors that plague most objective indicators of success. Perhaps the most serious potential error in production scores is criterion contamination. Contamination occurs when job incumbents' production is portrayed so that some appear better than their individual contributions warrant, while others appear poorer. Contamination is likely to be present whenever situational factors affect criterion scores in a stable way. It is important to note that such biasing effects are not detectable from reliability data. Thus, in the context of the present study, high reliability of the production "scores" says nothing about possible criterion contamination in these scores.

In this study and in some others (Larriva, 1975; Brown, Wood, & Harris, 1975), raw production statistics have been adjusted in various ways to reduce contaminating influences. However, it may be very difficult to partial out all or even a large proportion of these biasing influences. For one thing, an adjustment that equalizes mean production across districts or stations assumes random assignment in terms of effectiveness to districts and stations. This is unlikely in an operational setting. Thus, perhaps the reason that many studies have failed to find substantial relationships between predictor measures and production or adjusted production relates to criterion measurement problems. Such difficulties would restrict the magnitude of any relationship between predictors and performance.

Of course, performance ratings are fraught with their own problems. These difficulties have been well-documented elsewhere (e.g., Campbell, Dunnette, Lawler, & Weick, 1970; Whister & Harper, 1962), but particular pitfalls may be present in some of the kinds of ratings used in studies reviewed in a previous section of this report. For example, ratings, rankings, or nominations may be assigned by a commanding officer primarily on a reputation rather than on good knowledge of recruiters' day-to-day job activities. Also, rating forms utilizing dimensions that are not well conceived or well defined may cause commanding officers or others providing the criterion ratings to make imprecise evaluations of recruiters.

On the other hand, careful attention paid to defining performance dimensions and to selecting the proper persons to provide ratings may make "subjective" performance evaluations reasonable estimates of recruiters' true effectiveness, uncontaminated by situational constraints. Referring to Dunnette's (1963) modified model for test validation and selection research (see Figure 2) shows that in the present study predictor measures are probably being evaluated against job behaviors as rated by supervisors and recruiters. That is, with the behavior-oriented rating scales especially, raters are asked to identify and to evaluate behaviors that recruiters exhibit in performing on the job. Such judgments are a step removed from the consequences in Dunnette's model. Because they are evaluations of behavior rather than of consequences, such ratings may avoid situational influences on performance and thus should be, in this sense, better criteria against which to validate predictor measures.

Using a different conceptual framework to make a similar argument, the ratings gathered in the present study reflect performance more than effectiveness (Campbell, Dunnette, Arvey, & Hellervik, 1973). Campbell et al. define performance as job behavior that can be evaluated in terms of its contribution to organizational goals. Effectiveness, as they define it, refers to organizational outcomes (i.e., "consequences" in Figure 2) for which an individual may be only partially responsible. Again, for test validation purposes, the measurement of performance (i.e., behavior that can be evaluated, rather than outcome indicators such as those Campbell et al. term "effectiveness") appears most conceptually reasonable.

For these reasons, carefully developed rating scales and well-placed, well-prepared raters may provide the best estimates of military recruiter performance. In the present study, self, peer, and supervisory ratings proved to be reasonably reliable and, as criteria, they provided considerably higher validities than an adjusted production index did.

Relationships Between Predictor Composites and Recruiter Performance: Some Perspective

Estimated cross-validities for the predictor composites developed for Sales, Human Relations, and Administrative Skills and for Overall Performance are all significantly different from zero at the .01 level for the Navy sample and at the .05 level for the Marine Corps sample. Also, the methods used for estimating these cross-validities are conservative;

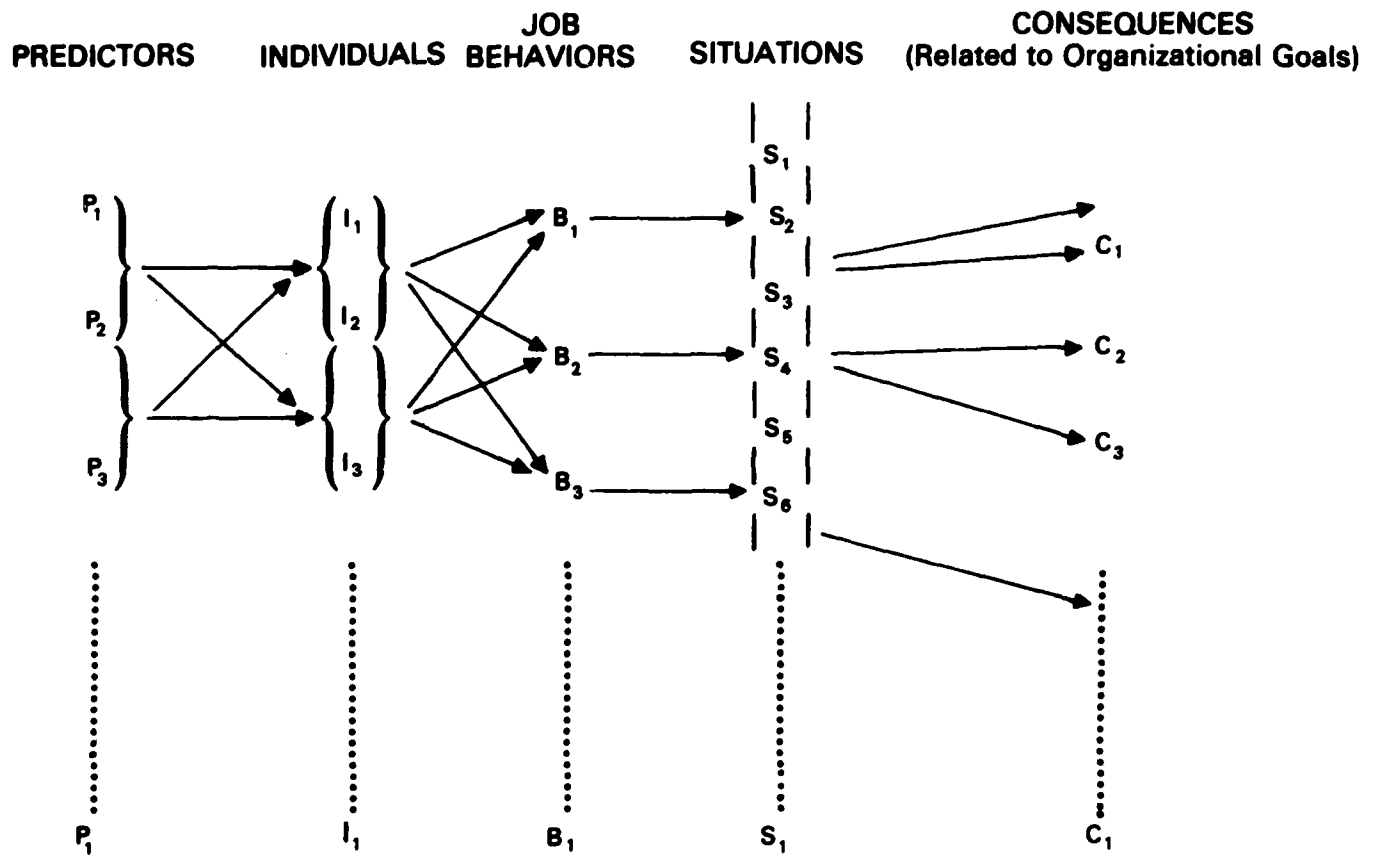


Figure 2. Dunnette's modified model for test validation and selection research.

the actual validities may be higher. Moreover, the pattern of relationships between predictors and criteria is conceptually reasonable, suggesting stable, systematic linkages between these predictor composites and recruiter performance. Nevertheless, one must still ask how useful the predictor composites are likely to be in practice.

One way of looking at the utility of the predictor composites is to employ the Taylor-Russell Tables (Taylor & Russell, 1939) to assess projected gains in the levels of performance expected from individuals selected according to the new set of predictors. Given a selection ratio of .50 (i.e., half of the persons applying for or being considered for recruiting duty actually become recruiters), a validity coefficient of .20 means that 56 percent of the persons selected for recruiting duty can be expected to be above average using the present levels of recruiter effectiveness as a standard. A validity coefficient of .25 suggests that 58 percent of the persons selected using the predictor battery can be expected to achieve an above-average performance. The corresponding figures for validity coefficients of .30, .35, and .40 are 60, 61, and 63 percent respectively. Thus, given the magnitude of the validity obtained for the predictor composites, and remembering that these validities are probably underestimates, it can be projected that the percent increase in recruiters performing above the present average would be 5 to 12 percent (depending on the performance criterion) if the recommended predictor composites are used to select individuals for recruiting duty.

Comparison of 16PF and Predictor Composite Validity Results

Since validity results were obtained in this study for both the 16PF (see Appendix A) and a newly developed predictor battery, the question arises as to which of the two screening devices performed better. Unfortunately, results of this study cannot answer that question conclusively.

First, the administration of the two predictor measures occurred at different times in the careers of persons in our sample. The 16PF was administered to members of our Navy sample either before selection as recruiters or after selection but during the training program before they went on recruiting duty. The predictors developed in the present study were of course administered to incumbents who had served some time on recruiting duty. Thus, data for the 16PF validity analysis conform more closely to the predictive validity model than do data for the predictor composite analysis. This means that validity coefficients associated with the 16PF may provide more faithful estimates of the validity of real interest than is provided by the predictor composite concurrent validity study results.

On the other hand, neither the predictor composites developed in the present study nor the performance measures correlated highly with tenure. Therefore, criterion-predictor composite relationships are not unduly affected by amount of time on the job, which in turn indicates that concurrent cross-validity estimates may be good estimates of the composites' predictive ability.

Still, it is obviously difficult to compare directly the results of the two analyses. The general magnitude of correlations with performance criteria is higher for the predictor composites than for the 16PF composite, especially for the Human Relations Skills, Administrative Skills, and Overall Performance criteria. However, the 16PF composite does correlate .21 with Selling Skills and .19 with Production, and several individual 16PF scales correlate in the .20s with performance criteria. These results suggest at least a modicum of validity (see Appendix A for these results). Ideally, both sets of measures should be studied in a predictive validity framework and the relationships between these measures and performance examined to assess their validity. At present, it can only be concluded that the 16PF possesses some validity for predicting Selling Skills, Overall Performance, and Production, and that our predictor battery shows good promise for predicting Navy recruiter success, most notably in the Selling and Administrative Skills areas. The predictor battery developed in this study shows even more promise for predicting Marine Corps recruiter effectiveness, particularly for the Administrative Skills and Overall Performance criteria.

RECOMMENDATIONS

Studying the Predictive Validity of the Predictor Composite

Cross-validity coefficients computed in the present study are only estimates of the composites' predictive validity. These estimates appear to be reasonable, even conservative, but the only way to ensure that they are is to perform a predictive study. Thus, paper-and-pencil measures consisting of those parts of the predictor battery found to be valid in the present study should be administered to a sample of enlisted personnel who are being considered for recruiting duty. (See page 54 for a more ideal approach to developing an improved predictor battery.) These persons should be led to believe that scores on the battery might be used for selection purposes, although, according to this plan, scores on the battery would not affect selection decisions. Presumably, some of the persons taking the battery would subsequently be placed in recruiting billets. After 6 months to a year on recruiting duty, performance information could be gathered for those persons, and the relationship between their performance and scores on the predictor composites evaluated.

As a head start on such a project, the Navy Recruiting Command is administering the battery containing the predictor composites to 200 persons being considered for recruiting assignments. Thus, with minimal additional effort, merely continuing to administer the battery to appropriate Fleet personnel and gathering performance ratings after an adequate number of these persons have served in a recruiting billet for 6 to 12 months, the Navy can determine the predictive validity of the battery developed in this study.

Studying the "Fakability" of the Predictor Battery

The present study has suggested that several items and scales from the personality, interest, and biographical domains show reasonably high validity as indicators of recruiter performance. However, the administrative set for persons in our study may have been different from the kind of set they would have had if they were taking the inventories as applicants before being accepted for recruiting duty. Persons completing a test or inventory before they are hired may try to respond to items as they think the selection decision-makers might want them to, while the recruiters in our concurrent validity sample were less likely to respond in that manner because they had nothing to gain from looking good.

Fortunately, at least two research studies (Dunnette, McCartney, Carlson, & Kirchner, 1962; Abrahams, Neumann, & Githens, 1971) suggest that persons taking personality or interest inventories in an actual employment setting tend to respond much more honestly, i.e., fake less, than persons who are instructed to fake their responses. In fact, the Dunnette et al. study indicates that the mean scores of a group applying for employment are very similar to those of a group tested only for research.

Thus, responses made by our concurrent validity group (research set) will probably be similar to those of persons who are being considered for

recruiting duty. However, the only way to ensure that the predictors are not fakable in this sense is to perform an empirical test of the hypothesis. The predictor data referred to earlier can also be used to evaluate this kind of fakability. Mean responses of the concurrent validity sample can be compared to the mean scores obtained by the 200 persons taking the same inventories in a selection setting. Those items and scales showing good validity and small differences in scores between these two groups are, of course, the best candidates for a revised predictor battery.

One way to decrease the probability of faking predictor battery responses is to administer the battery to Navy or Marine Corps personnel well before the time they might be considered for recruiting duty. For example, upon becoming a third-class petty officer, individuals in the Navy might be routinely required to complete the predictor battery. Scores on the battery could be stored and used subsequently for selection purposes if the man or woman was being considered for a recruiter assignment.

Developing Additional Measures of Valid Constructs

Based on results of this study and the studies reviewed earlier in this report, certain constructs--e.g., traits, personal characteristics, and interest patterns--appear to be associated with military recruiter effectiveness. Additional measures, items and scales should be developed to tap the constructs that seem to be valid indicators of recruiter performance. Extending the coverage of these constructs should add to the reliability of measurement and to the validity of the battery.

REFERENCES

- Abrahams, N. M., Neumann, I., & Githens, W. H. Faking vocational interests: Simulated versus real life motivation. Personnel Psychology, 1971, 24, 5-12.
- Abrahams, N. M., Neumann, I., & Rimland, B. Preliminary validation of an interest inventory for selection of Navy recruiters (NPTRL Research Memorandum SRM-73-3). San Diego, CA: Navy Personnel and Training Research Laboratory, April 1973.
- Arima, J. K. Determinants and a measure of Navy recruiter effectiveness. Prepared for the Navy Personnel Research and Development Center, San Diego, CA, 1977.
- Bennett, J. T., & Haber, S. E. Selection, deployment, and evaluation of Marine Corps recruiters. Washington, D. C.: The George Washington University, School of Engineering and Applied Science, Institute of Management Science and Engineering, Project NR347020, Office of Naval Research, 1973.
- Borman, W. C. First-line supervisor validation study. Minneapolis, MN: Personnel Decisions, Inc., 1973.
- Borman, W. C., Dunnette, M. D., & Johnson, P. D. The development and evaluation of a behavior based naval officer performance assessment package. Minneapolis, MN: Personnel Decisions, Inc., 1974.
- Borman, W. C., Hough, L. M., & Dunnette, M. D. Development of behaviorally based rating scales for evaluating the performance of U. S. Navy recruiters (NPRDC Tech. Rep. 76-31). San Diego, CA: Navy Personnel Research and Development Center, February 1976.
- Bray, D. W., & Grant, D. L. The assesement center in the measurement of potential for business management. Psychological Monographs, 1966, 80(17).
- Brogden, H. E., & Taylor, E. K. The validity of recruiter selection instruments at various points of cut (Report No. 781). Washington, D. C.: Personnel Research Section, April 1949.
- Brown, G. H., Wood, M. D., & Harris, J. D. Army recruiters: Criterion development and preliminary validation of selection procedure (FR-ED-75-8). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, April 1975.
- Campbell, D. P. Strong-Campbell Interest Inventory. Stanford, CA: Stanford University Press, 1974.
- Campbell, J. P., Dunnette, M. D., Arvey, R. P., & Hellervik, L. W. The development and evaluation of behaviorally based rating scales. Journal of Applied Psychology, 1973, 57, 15-22.

- Campbell, J. P., Dunnette, M. D., Lawler, E. E., & Weick, K. E. Managerial behavior, performance, and effectiveness. New York, NY: McGraw-Hill, 1970.
- Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. Handbook for the Sixteen Personality Factor Questionnaire (16PF). Champaign, IL: Institute for Personality and Ability Testing, 1970.
- Dunnette, M. D. A modified model for test validation and selection research. Journal of Applied Psychology, 1963, 47, 317-323.
- Dunnette, M. D. Multiple assessment procedures in identifying and developing managerial talent. In P. McReynolds (Ed.), Advances in psychological assessment, Vol. II. Palo Alto, CA: Science and Behavior Books, 1971.
- Dunnette, M. D. Aptitudes, abilities, and skills. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago, IL: Rand McNally, 1976 (Chapter 11). (a)
- Dunnette, M. D. Validation of selection tests and inventories for Federated Insurance Company administrative and clerical jobs. Minneapolis, MN: Personnel Decisions, Inc., December 1976. (b)
- Dunnette, M. D. Sales effectiveness scale. Minneapolis, MN: Personnel Decisions Research Institute, 1976.
- Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. A study of faking behavior on a forced-choice self-description checklist. Personnel Psychology, 1962, 15, 13-24.
- Dunnette, M. D., & Motowidlo, S. J. Development of a personnel selection and career assessment system for police officers for patrol, investigative, supervisory, and command positions. Minneapolis: Personnel Decisions, Inc., June 1975.
- Federal Executive Agencies. Uniform guidelines on employee selection procedures. Federal Register, 1976, 41(136), 29016-29022.
- Ghiselli, E. E. The forced-choice technique in self-description. Personnel Psychology, 1954, 7, 201-208.
- Ghiselli, E. E. Theory of psychological measurement. New York, NY: McGraw-Hill, 1964.
- Gough, H. G. California psychological inventory. Palo Alto, CA: Consulting Psychologists Press, Inc., 1969.
- Graf, R. G., & Brower, D. B. The development of an interest inventory for the selection of Marine Corps recruiters. San Diego, CA: Navy Personnel Research and Development Center, June 1976.
- Jackson, D. N. Personality research form. Coshen, NY: Research Psychologists Press, Inc., 1974.

Krug, S. E. Psychological and demographic predictors of success as a naval recruiter (Final Report). Washington, D. C.: Navy Recruiting Command, November 1972.

Larriba, R. F. U. S. Marine Corps recruiter performance prediction study. Unpublished manuscript, 1975.

Massey, I. H., & Mullins, C. J. Validation of the recruiter-salesman selection test. Lackland Air Force Base, TX: Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command, February 1966.

Rosse, R. L., & Hellervik, L. W. Validation study: Predicting department manager effectiveness. Minneapolis, MN: Personnel Decisions, Inc., September 1975.

Schwab, D. P., Henneman, H. G., III, & DeCotiis, T. A. Behaviorally anchored rating scales: A review of the literature. Personnel Psychology, 1975, 28, 549-562.

Taylor, H. E., & Russell, J. T. The relationships of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. Journal of Applied Psychology, 1939, 23, 565-578.

Tellegen, A. Differential personality questionnaire. Minneapolis, MN: University of Minnesota, 1976.

Whisler, T. L., & Harper, S. F (Eds.). Performance appraisal: Research and practice. Chicago, IL: The University of Chicago, 1962 (Holt, Rinehart, and Winston).

Wollack, L., & Kipnis, D. Development of a device for selecting recruiters (Task assignment PF-016-05-003-W2). Washington, D. C.: U. S. Naval Personnel Research Field Activity, March 1960.

APPENDIX A
16PF VALIDITY STUDY

16PF VALIDITY STUDY

Procedures Related to Validating the 16PF

The 16PF is a personality inventory developed by Cattell (Cattell, Eber, & Tatsuoka, 1970). It has been used in a variety of applied settings (e.g., personnel selection, occupational placement, vocational guidance, clinical diagnoses, assessment of accident proneness, and school performance prediction). The inventory yields scores on the following 16 dichotomous personality traits:

Factor	Low Score Description	High Score Description
1	Reserved, detached, critical, aloof, stiff	Outgoing, warmhearted, easy-going, participating
2	Dull, low intelligence	Bright, high intelligence
3	Affected by feelings, emotionally less stable, easily upset, changeable	Emotionally stable, mature, faces reality, calm
4	Humble, mild, easily led, docile, accommodating	Assertive, aggressive, competitive, stubborn
5	Sober, taciturn, serious	Happy-go-lucky, enthusiastic
6	Expedient, disregards rules	Conscientious, persistent, moralistic, staid
7	Shy, timid, threat-sensitive	Venturesome, uninhibited, socially bold
8	Tough-minded, self-reliant, realistic	Tender-minded, sensitive, clinging, overprotected
9	Trusting, accepting conditions	Suspicious, hard to fool
10	Practical, "down-to-earth"	Imaginative, bohemian, absent-minded
11	Forthright, unpretentious, genuine but socially clumsy	Astute, polished, socially aware
12	Self-assured, placid, secure, complacent, serene	Apprehensive, self-reproaching, insecure, worrying, troubled
13	Conservative, respecting traditional ideas	Experimenting, liberal, free-thinking
14	Group dependent, a "joiner" and sound follower	Self-sufficient, resourceful, prefers own decisions
15	Undisciplined self-conflict, lax, follows own urges, careless of social rules	Controlled, exacting will power, socially precise, compulsive, following self-image
16	Relaxed, tranquil, torpid, unfrustrated, composed	Tense, frustrated, driven, overwrought

Scores on the 16PF scale were available for Navy recruiters who had taken this inventory either before being selected for recruiting duty or during their training experience after selection. Also available were composite scores formed by applying a prediction equation used by the Navy Recruiting Command (Krug, 1972). The prediction equation employed empirically derived weights for four of the 16PF scales and several other variables. Thus, as a spinoff of this study 16PF scale scores and the composite scores were correlated with performance criteria for the 191 Navy recruiters who had completed the 16PF and supplemental questions previously.

16PF Validation Results

Correlations between 16PF scale scores and the five performance criteria appear in Table A-1. Results show that six of the 16 scales are related to Selling Skills with correlations of .20 or greater. Only two scales correlate .20 or greater with Human Relations Skills and four correlate in the .20s with Overall Performance. No 16PF scales correlate as highly as .20 with Administrative Skills or Production.

Relationships between the composite and each criterion also appear in the table. The composite correlates .21 with Selling Skills and .19 with Production. Correlations between the other three criteria and the 16PF composite are lower.

Table A-1
Correlations Between 16PF Scale Scores
and the Five Criteria (N = 191)

16PF Scales	Selling Skills	Human Relations Skills	Adminis- trative Skills	Overall Perform- ance	Produc- tivity
1	.22	.18	.02	.14	.03
2	.08	.14	.17	.15	.08
3	.01	.09	.00	.05	-.01
4	.20	.12	.11	.20	-.04
5	.27	.21	.02	.23	-.13
6	-.08	.04	.12	.03	-.11
7	.21	.18	.12	.23	-.07
8	-.22	.09	.06	-.14	.00
9	.19	.08	-.03	.14	.13
10	.09	.01	.06	.08	.16
11	-.06	-.05	.13	.05	-.01
12	.02	.04	-.01	-.02	.02
13	.19	.09	-.08	.11	-.08
14	-.20	-.20	-.05	-.24	-.05
15	-.06	.06	.15	.04	.05
16	-.05	-.12	-.06	-.11	-.02
Composite Score ^a	.21	.04	.15	.15	.19

^aSee Krug (1972) for a description of this composite, formed by empirically keying the 16PF and several other variables.

APPENDIX B

SCALE DEVELOPMENT STRATEGIES AND DEFINITIONS OF RATING DIMENSIONS

SCALE DEVELOPMENT STRATEGIES AND DEFINITIONS OF RATING DIMENSIONS

1. Trait dimensions. A literature review of selection studies directed at sales jobs and interviews with several Navy recruiters suggested several traits that appeared to be important for performing effectively as a Navy recruiter. An attempt was made to summarize the trait list by combining traits or eliminating traits when two or more of them were close in meaning. These procedures resulted in eight trait dimensions: Initiative, Judgment, Imagination and Originality, Personal Impact, Confidence, Achievement Orientation, Organization, and Warmth.

2. Behavior dimensions. In a previous study (Borman, Hough, & Dunnette, 1976) 40 Navy recruiters and 21 of their immediate supervisors provided over 1000 behavior examples of effective, adequate, and ineffective recruiter performance. These examples were content-analyzed by three psychologists, the process yielding nine dimensions of recruiter performance. The behavior examples were subsequently retranslated by an independent sample of 77 recruiters using methods outlined by Smith and Kendall (1963). Eight behaviorally anchored dimensions survived the retranslation process; one dimension, salesmanship skills, was dropped because in categorizing examples confusion arose between that dimension and several others. The final seven dimensions were Locating and Contacting Qualified Prospects, Gaining and Maintaining Rapport, Obtaining Information from Prospects and Making Good Person-Navy Fits, Establishing and Maintaining Good Relationships in the Community, Providing Knowledgeable and Accurate Information about the Navy, Administrative Skills, and Supporting Other Recruiters and the Command.

3. Multidimensional scaling dimensions. A similarity judgment task was developed to discover the underlying dimensionality of the content of behavior examples written for the Navy recruiter job. Sixty behavior examples were selected randomly from all of those rated as reflecting effective performance, and all possible pairs of these examples (1770 in all) were generated and randomly assigned without replacement to 20 different protocols so that each protocol contained approximately 89 different item pairs. Thus, each of the 1770 possible pairs of behavior examples appeared once in one of these 20 protocols. The same procedure was used to develop four more sets of 20 protocols, resulting in each item pair appearing exactly five times in 100 protocols. In addition, for each protocol, 14 or 15 item pairs were randomly selected (again without replacement) and repeated within each protocol in such a way that they appeared at least 20 items apart from the same item pair. These repeat item pairs allowed for an estimate of within-rater reliability.

Each item pair in the protocols required the respondent to assess the similarity of the job facets represented by the two behavior examples. A 4-point rating scale was used with "4" indicating that the two examples represented completely different facets of the recruiter job and "1" indicating that the two examples represented identical or very similar facets of the job.

One hundred Navy recruiters completed the protocols, with 16 of them performing the task unreliably ($r < .50$ for the correlation between the 14 or 15 pairs of similarity ratings). Other recruiters responded to these protocols until every one was reliably completed.

The consistency in responses across recruiter judges was then assessed by an analysis of variance (ANOVA) comparing the variance of different judges' similarity ratings of the same item pairs with the total variance across all similarity ratings. Such an ANOVA procedure was possible because five independent estimates of the similarity of each of 1770 item pairs were available. An intraclass correlation coefficient that summarizes this across-judge agreement was .71 ($p < .001$, $df = 1770, 7080$), indicating that judges agreed closely in their similarity ratings of item pairs. Therefore, mean similarity ratings were computed for each of the 1770 item pairs and used to form a 60 x 60 similarity matrix.

The similarity matrix was then submitted to a nonmetric multidimensional scaling analysis (Shepard, 1962; Kruskal, 1964), with each solution rotated to the varimax criterion. Stress values and interpretability of dimensions were used to determine the preferred solution. The five-dimension solution provided eight interpretable job dimensions because three of the dimensions were bipolar. The resultant eight scales were Prospecting, Gathering Information about Applicants, Planning and Organizing Own Recruiting Schedule and Practices, Expending Extra Effort to Aid Applicants or Recruits, Preparing Recruits and their Parents for Navy Life, Salesmanship, Developing Productive Relationships in the Community, and Gaining and Maintaining Friendly Relationships with Prospects.

4. Factor analysis dimensions. The eight behavior dimensions were used in a previous study (Borman, Hough, & Dunnette, 1976) to rate the performance of 24 Navy recruiters stationed in the Minneapolis-St. Paul district. Supervisory, peer, and self ratings were gathered, the ratings were pooled across the three sources, and the pooled ratings on the eight dimensions were intercorrelated. A principal-components factor analysis of these intercorrelations followed by a varimax rotation of the three-factor solution revealed three conceptually meaningful dimensions of recruiter performance. The factors and the proportion of common variance accounted for by each appear below.

- I. Prospecting and Selling (36%)
- II. Planning, Organizing, and Administrative Skills (37%)
- III. Developing Good Navy-Community Relations and Expending Extra Effort to Aid Prospects and Recruits (27%)

RATING DIMENSIONS AND DEFINITIONS OF PRETEST RATING SCALES

1. Initiative. Displaying a willingness to seek out and to accept responsibility; working well on one's own.
2. Judgment. Ability to handle day-to-day problems and situations in a practical and efficient way; using information wisely in making decisions.
3. Imagination and Originality. Being resourceful and creative in determining productive courses of action.
4. Personal Impact. Charisma; creating a positive first impression.
5. Confidence. Projecting self-assurance and decisiveness in dealing with others.
6. Achievement Orientation. Enjoying involvement in the process of achieving significant goals; capacity to stick with a task and to expend considerable physical and mental energy to attain an objective.
7. Organization. Planning ahead effectively; being well-organized and efficient in performing tasks and duties.
8. Warmth. Relating easily and naturally to others; being friendly, outgoing, and warm.
9. Locating and Contacting Qualified Prospects. Prospecting effectively; contacting large numbers of persons likely to join the Navy; getting prospects into the office.

Skillfully using the telephone referrals, recruits, advertising ideas, and special events to contact young persons eligible for Navy service and get their attention.

Knowing where and when to prospect; ability to persist in prospecting and following up on leads even under considerable adversity.
10. Gaining and Maintaining Rapport. Being hospitable to prospects in the office.

Gaining the trust and respect of prospects.

Adjusting to prospects' styles and acting appropriately with different types of prospects.
11. Salesmanship Skills. Skillfully persuading prospects to join the Navy; using Navy benefits and opportunities effectively to sell the Navy.

Closing skills; adapting selling techniques appropriately to different prospects.

Effectively overcoming objections to joining the Navy.

12. Obtaining Information from Prospects and Making Good Person-Navy Fits. Exercising listening skills; making accurate judgments about prospects' needs, programs desired, and the like, based on good interviewing skills.

Effectively obtaining information about prospects from other sources (e.g., high school principal) to assess their qualifications and needs.

Assessing accurately prospects' eligibility for various programs.

13. Establishing and Maintaining Good Relationships in the Community. Contacting and working effectively with high school counselors, newspaper editors, radio and TV personnel, and others capable of helping recruiters to enlist prospects.

Building a good reputation for the Navy by developing positive relationships with persons in the community; presenting a good Navy image in the community.

Establishing and maintaining good relationships with families of prospects.

14. Providing Knowledgeable and Accurate Information about the Navy. Displaying considerable knowledge about Navy programs, schools, and educational opportunities; answering questions about the Navy in a competent manner.

Providing accurate information about Navy life; skillfully relaying information about boot camp so that the prospective recruit is informed about what to expect but not discouraged from joining the Navy.

Being up to date on Recruiting Manual changes and on other directions pertaining to program or school changes.

15. Administrative Skills. Planning ahead; organizing time efficiently. Completing paperwork accurately and on time; keeping track of appointments.

Not wasting time.

16. Supporting Other Recruiters and the Command. Coordinating activities with other recruiters to maximize the productivity of the station and district.

Using own skills and time to support other Navy recruiters when appropriate; providing constructive feedback to other Navy recruiters concerning their skills and style; providing helpful tips to new recruiters.

Pitching in to support orders and directives from higher levels.

17. Prospecting. Using various prospecting techniques to get qualified young persons interested in joining the Navy.

Effectively using other persons to provide referrals or to help prospect for potential recruits.

18. Gathering Information about Applicants. Checking up on an applicant's background.

Asking appropriate persons about applicants in order to gather information by which to accept or reject them.

19. Planning and Organizing Own Recruiting Schedule and Practices. Establishing a realistic and efficient recruiting time schedule to meet goals.

Effectively analyzing own recruiting techniques.

20. Expending Extra Effort to Aid Applicants or Recruits. Willingly offering and providing assistance to recruits or applicants who have a problem or need.

Helping other recruiters to meet a recruit's or applicant's special needs.

21. Preparing Recruits and Their Parents for Navy Life. Fully informing recruits and their parents about boot camp and Navy life.

Developing good relationships with the families of recruits.

22. Salesmanship. Using knowledge of Navy jobs and the qualifications necessary for those jobs to sell the Navy.

Listening to the prospect and then making an appropriate and effective sales pitch.

23. Developing Productive Relationships in the Community. Volunteering for community events or projects in order to enhance the image of the Navy and Marine Corps image.

Developing relationships with recruits and with persons in the community as sources of referrals.

24. Gaining and Maintaining Friendly Relationships with Prospects. Being friendly, outgoing, and sincere with all prospects regardless of their backgrounds.

Adapting recruiting style to the individual prospect.

25. Prospecting Skills. Effectively contacting qualified prospects and interesting them in the Navy.

Locating potential recruits through other persons who provide referrals, free advertising and the like.

Knowing how, when, and where to use prospecting resources such as the telephone, referrals, recruits, and advertising ideas to get the attention of young persons qualified for Navy service.

26. Selling Skills. Skillfully using knowledge of Navy benefits, programs, schools, and educational opportunities to sell the Navy to prospects.

Effectively answering prospects' questions about the Navy and overcoming their objections to entering the naval service.

Using selling techniques that get qualified prospects into the Navy.

27. Human Relations Skills. Using good listening and interviewing skills with prospects.

Effectively adjusting to different types of prospects and gaining their trust.

Providing support to the station, zone, district, and to other recruiters by attending more to organizational quotas than personal quotas and by working to enhance team effort with fellow recruiters.

28. Likability.¹ Overall subjective judgment of how well you like this individual.

29. Overall Effectiveness as a Recruiter. Overall level of effectiveness displayed in performing as a Navy or Marine Corps recruiter.

¹The Likability dimension was used for special research purposes only and was not included as a performance criterion measure in any subsequent analyses.

REFERENCES

- Borman, W. C., Hough, L. M., & Dunnette, M. D. Development of behaviorally based rating scales for evaluating the performance of U. S. Navy recruiters. Minneapolis, MN: Personnel Decisions Research Institute, 1976.
- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964, 29, 1-24.
- Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. Psychometrika, 1962, 27, 125-140.
- Smith, P. C., & Kendall, L. M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 1963, 47, 149-155.

APPENDIX C

MAIN SAMPLE PERFORMANCE RATING SCALE PACKAGE

(For Navy only; the Marine Corps package was the same except for reference to the "Marine Corps" rather than the "Navy.")

NAVY RECRUITER PERFORMANCE CATEGORIES

Form A

I. INTRODUCTION

Eight categories or facets of job performance for the Navy Recruiter are named and defined on the following eight pages. The eight performance categories are:

- A. LOCATING AND CONTACTING QUALIFIED PROSPECTS
- B. GAINING AND MAINTAINING RAPPORT
- C. OBTAINING INFORMATION FROM PROSPECTS AND MAKING GOOD PERSON-NAVY FITS
- D. SALESMANSHIP SKILLS
- E. ESTABLISHING AND MAINTAINING GOOD RELATIONSHIPS IN THE COMMUNITY
- F. PROVIDING KNOWLEDGEABLE AND ACCURATE INFORMATION ABOUT THE NAVY
- G. ADMINISTRATIVE SKILLS
- H. SUPPORTING OTHER RECRUITERS AND THE COMMAND

First, notice that each of the eight Performance Categories is labeled and defined carefully at the top of the page.

Then, four levels of job performance effectiveness are listed and three descriptive statements are provided to help you define or get a fix on each level. The four levels of effectiveness are:

9 or 10: EXTREMELY EFFECTIVE

6, 7, or 8: EFFECTIVE

3, 4, or 5: MARGINAL

1 or 2: INEFFECTIVE

The following are guidelines to help you make accurate ratings of each Navy Recruiter's job performance. You should record your actual ratings in the booklet titled, "Navy Recruiter Rating Scales."

The most important part of the entire rating task is for you to read all the descriptive statements very thoroughly so that you have firm knowledge of the kind of performance that defines each of the above four levels for each Performance Category.

II. GUIDELINES FOR MAKING NAVY RECRUITER JOB PERFORMANCE RATINGS

1. In rating the job performance of a specific Navy recruiter, first read the category description at the top of the page to get a general idea of what this dimension means.

typical performance is really slightly better than the level of performance given by the three descriptive statements corresponding to MARGINAL, even though it definitely does not stack up to the level portrayed by statements at the EFFECTIVE level. You would then give this recruiter a rating of "5."

2. Next, read the set of descriptive statements for each effectiveness level and decide which set of behaviors most closely matches that recruiter's typical performance for this Performance Category. Does he/she typically exhibit behavior that's Extremely Effective, Effective, Marginal, or Ineffective.

On the other hand, if the recruiter were judged to be squarely in the middle of the MARGINAL range, again based upon a match between the descriptive statements and his or her typical performance, a rating of "4" would be appropriate. In other words, use the two or three numbers within each effectiveness level to make finer distinctions about recruiters rated within each level.

3. After the above decision has been made, consider how this recruiter stacks up compared to others who are at the same level of effectiveness. Ask yourself the following question: "Within this level of performance effectiveness, is this recruiter among the more effective or among the less effective?" For example, you might rate a recruiter as MARGINAL (3, 4, or 5) but feel that his or her

Please work through your ratings category by category. That is, evaluate all recruiters on the first Performance Category (Locating and Contacting Qualified Prospects) before moving on to Category B. Follow this procedure for all eight Performance Categories.

III. THINGS TO GUARD AGAINST

Several sources of error can contribute to inaccuracies in your ratings. Here are a few suggestions for overcoming them:

1. Consider each Performance Category separately from all the rest. An almost universal error in ratings is called HALO ERROR. It occurs when the rater gives about the same ratings to a person on all aspects of performance. Usually this error occurs because a rater has not taken enough time to get clearly in mind what each separate category of performance refers to. Remember we are asking you to describe or evaluate each ratee on eight different categories of performance. As you consider each of the persons you are rating, try to avoid getting into the habit of giving about the same rating to him/her on each performance category. Consider each category separately from all others. Be sure to rate all ratees on each category before going on to the next category.
2. Consider each ratee's performance over time and not on just one or two occasions. Another type of error occurs when a ratee is influenced by just a single event or a recent occurrence. As you consider each ratee's performance, think back over all the time you have known him/her and try to avoid being influenced by just one or two events. Base your ratings on all your observations of the person's performance and not just a few.
3. Avoid being misled by such things as appearance, education, family background, and other personal characteristics. Another common error in rating is called STEREOTYPE ERROR. It occurs when a rater allows himself/herself to be influenced by other things than what the person has actually done on the job. In considering each ratee's job performance, try to ignore everything else you may know about that person. Base your rating strictly on what the individual has done on the job.
4. Avoid using your own definitions for the various performance categories. A common reason for inaccurate ratings is that raters have different definitions of the performance categories. Terms such as "Gaining and Maintaining Rapport" and "Salesmanship Skills" can have different meanings for different raters. This is why it is so very important for you to read the definitions and behavior statements carefully for these performance categories. Avoid any previous impressions of what these things have meant to you. Base your ratings on the definitions which are provided in the rating booklet you have received.

A. LOCATING AND CONTACTING QUALIFIED PROSPECTS

Is adept at using a variety of prospecting tools effectively: prospect cards; radio, TV, and newspaper advertising; school contacts; personal appearances; telephone.

9 or 10

Extremely Effective Performance

Makes very judicious use of persons recruited recently or placed in CACHE to get names of qualified young men and women likely to join the Navy.

Is adept at using a variety of prospecting tools effectively: prospect cards; radio, TV, and newspaper advertising; school contacts; personal appearances; telephone.

6, 7, or 8

Effective Performance

Usually persuades prospects to visit the recruiting office by using the telephone effectively or by talking to prospects in person.

Follows up promising leads on potential recruits.

3, 4, or 5

Marginal Performance

Exerts effort prospecting, but does not seek recruits in optimal locations.

Tends to spend too much time with a prospect at the expense of contacting other prospects, or may spend time with a person who is disinterested or has something else to do.

1 or 2

Ineffective Performance

Uses very few prospecting tools; waits for prospects to come to him; almost always fails to follow up on leads, even promising ones.

Falls to become familiar with his recruiting area, and thus lacks knowledge about sources of recruits.

B. GAINING AND MAINTAINING RAPPORT

Being hospitable to prospects in the office; gaining the trust and respect of prospects; adjusting to applicants' styles and acting appropriately with different types of applicants.

9 or 10

Extremely Effective Performance

Deals very effectively with persons of all races; greets all prospects appropriately; is adept at setting them at ease and getting them to talk, regardless of their background, race, or personality.

Is adaptable, but not phony in interacting with all types of prospects; maintains a sincere, courteous, and friendly atmosphere in the office.

Answers prospects' questions politely and patiently, no matter how unimportant they seem.

6, 7, or 8

Effective Performance

Is almost always able to put prospects at ease when they first enter the office.

Expresses concern toward recruits and shows interest in their recruitment activities; for example, warmly wishes a recruit good luck in boot camp.

Shows interest in most persons who enter the office and interacts with them in a warm and friendly way.

3, 4, or 5

Marginal Performance

Has a standard approach with all persons which, at times, is inappropriate, such as a manner of greeting, speech, or telling sea stories.

Occasionally appears disinterested when with a prospect; sometimes forgets an applicant's name.

Is discourteous at times; for example, will sometimes interrupt an applicant while he/she is speaking.

1 or 2

Ineffective Performance

Ignores or is rude to applicants who do not seem, at first sight, to be "good Navy recruits."

Will interrupt an ongoing interview to interview another person or will perform other duties during an interview and give an impression of being disinterested in the person.

Is cold and impolite upon initial meeting and answers questions in a disinterested, nonpersonable way.

C. OBTAINING INFORMATION FROM PROSPECTS AND MAKING GOOD PERSON-NAVY FITS

Listening skills; making accurate judgments about prospects' needs, programs desired, etc., based on good interviewing skills; effectively obtaining information about prospects from other sources (e.g., high school principal) to assess their qualifications and needs; assessing accurately prospects' eligibility for various programs.

9 or 10

Extremely Effective Performance

Always listens attentively to applicant and then asks the right questions at the right time so that applicant reveals both positive and negative information and his/her interests and capabilities.

Always talks with prospects' school counselors, teachers, friends, and parents in a discreet, genuinely interested manner and thus obtains relevant information about prospects' interests, abilities, and when applicable, needs for waivers.

Consistently remembers and analyzes information gathered from prospects to make wise suggestions about programs to consider; very rarely mismatches an applicant's interests and abilities with Navy program requirements.

6, 7, or 8

Effective Performance

Listens to individuals' reasons for considering the Navy and asks questions about their interests.

Consistently makes routine checks on applicants to determine their eligibility for enlistment.

Tries to identify Navy programs that are well suited for the applicant's needs, interests, and capabilities.

3, 4, or 5

Marginal Performance

Sometimes talks too much without letting prospects or applicant talk enough; doesn't always learn enough about the individual to recommend appropriate programs.

May check a few references to learn about an applicant but does not thoroughly check them all.

Suggests programs prematurely or suggests programs that do not result in a good match for the individual and the Navy.

1 or 2

Ineffective Performance

Ignores applicants' stated interests; does not ask applicants what they are interested in.

Learns very little from others about applicants; checks very few, if any, references.

Suggests programs applicant is either not qualified for or not interested in.

D. SALESMANSHIP SKILLS

Skillfully persuading prospects to join the Navy; using Navy benefits and opportunities effectively to sell the Navy; closing skills; adapting selling techniques appropriately to different prospects; effectively overcoming objections to joining the Navy.

9 or 10

Extremely Effective Performance

Is able to sell almost any appropriate program to an applicant even when the applicant is set on a different program for which he/she does not qualify; describes Navy life in an appropriate and convincing way for each prospect; is exceptionally effective and quick in countering objections about the Navy or a program.

Knows just the right people to include or to mention while talking with a prospect; consistently identifies those benefits most likely to convince each prospect.

6, 7, or 8

Effective Performance

Is capable of selling several programs; shows films relevant to programs and to Navy life; may use cruise book effectively to help sell Navy.

Often mentions specific benefits likely to persuade a prospect; describes Navy life well; is prepared to counter frequently heard objections about the Navy.

3, 4, or 5

Marginal Performance

Describes Navy benefits, programs, and life in a way that is not always suited to the particular prospect.

Often ignores criticisms of Navy life a prospect makes.

1 or 2

Ineffective Performance

Oversells a specific program and loses prospects when they don't qualify for it; neglects to describe other important aspects of Navy life.

Is frequently unable to counter objections to joining the Navy; may end a conversation when prospect states an objection.

Frequently emphasizes benefits or aspects of Navy life which are irrelevant or unappealing to the particular prospect; may describe Navy life, benefits and programs in a bland, unenthusiastic manner.

E. ESTABLISHING AND MAINTAINING GOOD RELATIONSHIPS IN THE COMMUNITY

Contacting and working effectively with high school counselors, newspaper editors, radio and TV personnel, and others capable of helping recruiters to enlist prospects; building a good reputation for the Navy by developing positive relationships with persons in the community; establishing and maintaining good relationships with parents and family of prospects; presenting a good Navy image in the community.

9 or 10

Extremely Effective Performance

Is exceptionally adept at cultivating and maintaining excellent relationships with school counselors, teachers, principals, police, news media persons, local business persons, and other persons who are important for getting referrals and free advertising.

Is innovative in informing the public about the Navy; actively promotes the Navy and makes friends for the Navy while doing it; always distributes the most current Navy information.

Volunteers off-duty time to work on community projects, celebrations, parades, etc.

6, 7, or 8

Effective Performance

Spends productive time with individuals such as police, city government, or school officials; may lunch with them or distribute calendars, appointment books, or buttons to them and invite them for cocktails.

Arranges for interested persons such Navy activities as trips to the Naval Academy; keeps relevant persons informed of Navy activities.

Encourages principals, counselors, and other persons important to a prospect to call if they have any questions about the Navy.

3, 4, or 5

Marginal Performance

Contacts school officials only sporadically; keeps them waiting for information they want; relationships with counselors, teachers, and persons important to an applicant or recruit are distant and underdeveloped.

Is not alert to opportunities to promote the Navy; rarely volunteers off-duty time to promote the Navy and is unenthusiastic when approached to do something for the community; rarely accepts speaking invitations.

Is, at times, discourteous to persons in the community; for example, sends form letters to persons who have assisted him or other Navy recruiters; is not always alert to the family's desire for more information about the Navy and the program in which their son or daughter enlisted.

E. (Continued)

1 or 2

Ineffective Performance

Does not contact high school counselors; does not accept speaking engagements; drives around in car instead of getting out and meeting people.

Alienates persons in community or persons important to an applicant or recruit by ignoring them, not answering their questions, responding rudely, demanding information, or encouraging high school students to drop out of school; sometimes does not appear at recruiting presentations for which he/she is scheduled.

Presents negative image of the Navy by doing things like driving while intoxicated or speeding and honking impatiently at other drivers; may express dislike for the Navy or recruiting.

F. PROVIDING KNOWLEDGEABLE AND ACCURATE INFORMATION ABOUT THE NAVY

Displaying considerable knowledge about Navy programs, schools, and educational opportunities; providing accurate information about Navy life; being up to date on Recruiting Manual changes and on other directives pertaining to program or school changes and eligibility for programs; skillfully relaying information about boot camp so that prospect/recruit is informed about what to expect but is not discouraged from joining the Navy; answering questions about the Navy in a competent manner.

9 or 10

Extremely Effective Performance

Is consistently accurate and honest in describing Navy life; thoroughly prepares recruits for boot camp and Navy life.

Consistently provides accurate information about Navy rates, education, etc.; when uncertain, immediately obtains the correct information; provides detailed and complete information when appropriate and when asked.

Studies and learns as much Navy information as possible without decreasing recruiting time.

6, 7, or 8

Effective Performance

Keeps current on the basics of all Navy programs and rates.

Prepares applicants for recruitment process; when appropriate, describes parts of boot camp and Navy life to recruits.

Competently answers most questions prospects and recruits ask.

3, 4, or 5

Marginal Performance

Mentions specific aspects of boot camp, rates, programs, benefits, and life in the Navy, but descriptions are often incomplete.

Sometimes provides inaccurate information about rates, programs, boot camp, etc.

Is not well informed or knowledgeable about many programs, rates, and benefits of the Navy.

1 or 2

Ineffective Performance

Frequently fails to prepare recruits for life in the Navy; fails to describe boot camp; describes aspects of the Navy irrelevant to a prospect.

Spends little time in learning about Navy programs, rates, etc., and so is unable to describe many programs for which an applicant is qualified.

Often provides inaccurate information, thus causing recruits to expect events which never occur.

G. ADMINISTRATIVE SKILLS

Planning ahead; organizing time efficiently; completing paperwork accurately and on time; keeping track of appointments; not wasting time.

9 or 10

Extremely Effective Performance

Maintains a complete and accurate calendar of relevant community, school, and Navy events and holidays; schedules meetings, interviews, and physicals wisely in order to use own and others' time most efficiently.

Finishes all paperwork accurately and on time; for example, when going on leave, completes all paperwork prior to departing.

Devises monthly, weekly, and daily plans, and follows them in order to achieve own and district goals; when changes in daily schedule occur, completes other necessary tasks.

6, 7, or 8

Effective Performance

Completes paperwork on time and with few errors; may collect information by phone to complete paperwork on time.

Keeps a record of appointments and applicant physicals.

Generally uses time efficiently; does paperwork at times when not likely to meet prospects; on out-of-town calls, often meets with others along the way.

3, 4, or 5

Marginal Performance

Completes paperwork either just barely in time, or not on time; errors in paperwork are common; some forms are not sent in again, delaying enlistment process.

Fails to use dead time wisely; for example, might spend an undue amount of time providing prospects and applicants with transportation.

Tends to schedule appointments, physicals, etc. without considering other events.

1 or 2

Ineffective Performance

Uses wrong forms; forms invariably contain errors; consistently forgets to complete all the necessary forms.

Does not maintain a record of appointments, and so may schedule two at the same time, or schedule appointments too closely together; misses appointments.

Does not know what to do with time, needs to be told what to do.

H. SUPPORTING OTHER RECRUITERS AND THE COMMAND

Coordinating activities with other recruiters to maximize the productivity of the station and district; using own skills and/or time to support other Navy recruiters when appropriate; pitching in to support orders and directives from higher levels; providing constructive feedback to other Navy recruiters concerning their skills, style, etc.; providing helpful tips to new recruiters.

9 or 10

Extremely Effective Performance

Consistently works with others to improve selling and interviewing skills; provides valuable feedback to other recruiters; is always concerned about other recruiters' personal and professional difficulties and always helps them if possible.

Is always enthusiastic and works to build group spirit when other recruiters are down; enhances the team spirit even when he/she has personal problems; very seldom complains.

Always shares information such as legal shortcuts or important contacts which can increase zone production or lessen technical red tape.

6, 7, or 8

Effective Performance

Places station and district quotas above personal quotas; for example, will encourage prospects to sign up during current month if station needs more enlistments, rather than having prospects wait until the next month when he/she personally needs more.

Occasionally works with other recruiters to improve their interviewing or selling techniques; shares with other recruiters' prospects that he/she has attracted.

Helps other recruiters with paperwork; answers their phone when they are out; may get transcripts or parental consent for other recruiters.

3, 4, or 5

Marginal Performance

Occasionally gives negative feedback to other recruiters, even when it isn't warranted; does not provide feedback to other recruiters when it is needed or they want it.

Complains about production quotas or having to work extra time to get a recruit ready to ship.

Does not assist another recruiter when help is needed; gives ambiguous responses to other recruiters' questions; does not help new recruiters learn their job.

H. (Continued)

1 or 2

Ineffective Performance

Is quick to take personal credit when station achieves goals and quick to blame others when the station does not; lets others carry the recruiting load.

Interferes with the group effort by not answering the phone, neglecting to give messages to persons for whom intended, not communicating and coordinating schedule with others, or tying up the vehicles so other recruiters miss appointments.

Is very temperamental and cannot get along with other recruiters; may inappropriately interrupt other recruiters' interviews.

NAVY RECRUITER RATING SCALE

(EXAMPLE)¹

¹A rating scale of this type was provided for each of the performance dimensions used in the main study. The summary definition at the top of the sheet was the only explanation provided for the supplemental dimensions.

A. LOCATING AND CONTACTING QUALIFIED PROSPECTS

"Prospecting" effectively; contacting large numbers of persons likely to join the Navy; skillfully using the telephone, referrals, recruits, advertising ideas, special events, etc., to contact and get the attention of young persons eligible for Navy service; knowing where and when to prospect; ability to persist in prospecting and following up on leads even under considerable adversity; getting prospects into the office.

	Ineffective Performance		Marginal Performance			Effective Performance			Extremely Effective Performance	
Recruiters	1	2	3	4	5	6	7	8	9	10
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

APPENDIX D

NAVY RECRUITER BIOGRAPHICAL AND OPINION SURVEY
(Trial Battery)

NAVY RECRUITER BIOGRAPHICAL AND OPINION SURVEY

Name _____

1. Your age is _____.
2. How many years have you served in the Navy? _____ years
3. How many months have you served at your present station? _____ months
4. How many months have you served within the district? _____ months

SECTION 1: In this section, circle the one alternative that describes you best.

1. Sex
 - a. Male
 - b. Female
2. Race
 - a. White
 - b. Black
 - c. Spanish-American
 - d. Oriental
 - e. Native American
 - f. Other
3. Height
 - a. 5'4" or shorter
 - b. 5'5" to 5'7"
 - c. 5'8" to 5'10"
 - d. 5'11" to 6'1"
 - e. 6'2" to 6'4"
 - f. 6'5" or taller
4. Weight
 - a. 120 lbs. or less
 - b. 121 lbs. to 140 lbs.
 - c. 141 lbs. to 160 lbs.
 - d. 161 lbs. to 180 lbs.
 - e. 181 lbs. to 200 lbs.
 - f. 201 lbs. to 220 lbs.
 - g. 221 lbs. to 240 lbs.
 - h. 241 lbs. or more
5. How far did you go in school?
 - a. 8th grade or less
 - b. 9 to 11 years
 - c. Graduated from high school
 - d. 1 to 4 years of college
 - e. Bachelor's degree
 - f. Other

6. What was your marital status at the time you applied to be a recruiter?
- a. Single.
 - b. Separated.
 - c. Divorced.
 - e. Widowed.
 - f. Remarried.
 - g. Married.
7. What was your father's education?
- a. Grade school or less.
 - b. Some high school.
 - c. Completed high school.
 - d. Some college.
 - e. Completed college.
 - f. A graduate degree (M.A., M.S., Ph.D., etc.).
8. What was your mother's education?
- a. Grade school or less.
 - b. Some high school.
 - c. Completed high school.
 - d. Some college.
 - e. Completed college.
 - f. A graduate degree (M.A., M.S., Ph.D., etc.).
9. How many different towns did you live in before entering the Navy?
- a. 1
 - b. 2
 - c. 3
 - d. 4
 - e. 5
 - f. 6
 - g. 7
 - h. 8 or more
10. In what section of town did your family live longest while you were growing up?
- a. An exclusive section of town.
 - b. In a good but not the best section.
 - c. Average section of town.
 - d. In one of the poorer sections of town.
 - e. In a rural area.

11. How many times did you change schools before you were 16 years of age?
(Other than by graduation)
- a. Never.
 - b. Once.
 - c. Twice.
 - d. Three times.
 - e. Four or more times.
12. What size high school did you attend?
- a. Fewer than 100 students.
 - b. 100-499 students.
 - c. 500-999 students.
 - d. 1000-2000 students.
 - e. More than 2000 students.
13. As a youngster, how often were you a leader in your group of friends?
- a. Always.
 - b. Frequently.
 - c. Occasionally.
 - d. Seldom or never.
 - e. Was not a member of a group.
 - f. Can't remember.
14. When you were a child, did you feel that you received adequate recognition from your teachers for your work in school?
- a. Almost always.
 - b. Usually, but not always.
 - c. In a moderate amount.
 - d. Sometimes, but not often.
 - e. Almost never.
15. It is typical for you to live, eat, and breathe your job (in jobs you held prior to recruiting duty).
- a. Definitely agree.
 - b. Probably agree.
 - c. Not sure.
 - d. Probably disagree.
 - e. Definitely disagree.
16. When you were a high school student, were you:
- a. One of the most popular students.
 - b. More popular than most students.
 - c. About as popular as most students.
 - d. Not quite as popular as most students.

17. How often have you done extra work for a job which was not really required of you?
- a. Almost every day.
 - b. Several times a week.
 - c. About once a week.
 - d. Once every few weeks.
 - e. About once a month or less.
18. In group discussion or meetings, do you generally:
- a. Take an active role in discussions.
 - b. Remain relatively inactive, but try to learn from the discussion.
 - c. Remain relatively inactive, offering comments only rarely.
 - d. Rarely or never attend meetings.
19. About the best indication of a man's worth is how well he does his job.
- a. Definitely agree.
 - b. Probably agree.
 - c. Not sure.
 - d. Probably disagree.
 - e. Definitely disagree.
20. In high school, when friends came to you with their personal problems, how likely were you to go out of your way to give them help or advice?
- a. Much more likely than most people.
 - b. More likely than most people.
 - c. About the same as other people.
 - d. Less likely than most people.
 - e. Much less likely than most people.
21. On the average how much sleep do you require to feel really good?
- a. Less than 5 hours.
 - b. 5 to 6-1/2 hours.
 - c. 6-1/2 to 7 hours.
 - d. 7 to 8 hours.
 - e. 8 to 9 hours.
 - f. More than 9 hours.
22. In high school did you:
- a. Have many friends.
 - b. Have a few close friends.
 - c. Have hardly any friends.
 - d. Have no one you could call a friend.
23. Have you ever done door-to-door (not store) selling?
- a. Yes.
 - b. No.

24. Do you make a list of things to do when you know you will have a busy day?
- a. Yes, always.
 - b. Yes, usually.
 - c. Yes, sometimes.
 - d. Yes, but only rarely.
 - e. No, never.
25. How well have you felt you were able to understand the feelings of others?
- a. Very well.
 - b. Pretty well.
 - c. Fairly well.
 - d. Not very well.
26. How often do you find that your first impression of a person is the right one?
- a. Always.
 - b. Often.
 - c. Occasionally.
 - d. Rarely.
 - e. Never.
27. How well do you like to be around other people?
- a. I enjoy being with others very much; only rarely do I like to be by myself.
 - b. I usually enjoy being around others, occasionally preferring to be by myself.
 - c. I like being around other people sometimes and at other times I like to be by myself.
 - d. I prefer being by myself and only occasionally enjoy being around other people.
28. I notice little things about a person or a situation that others overlook.
- a. This happens to me almost all the time.
 - b. This often happens to me.
 - c. This has happened to me several times, but I wouldn't say this is generally true of me.
 - d. This very seldom happens to me.
 - e. This never happens to me.
29. Some people easily become involved in a task while others seldom really "dig into" a task or job. How involved do you usually become in a task or job?
- a. I often have trouble sticking with it; other things almost always seem to come up to distract my attention.

- b. I sometimes become involved in a task or job that interests me, but if I'm not very interested, I seldom become involved.
- c. I often become heavily involved in a task or job.
- d. I almost always become engrossed in tasks or jobs.

30. Do you feel you are a good detail person?

- a. Definitely yes; I am very detail oriented and attend closely to "nitty-gritties" of a task or job.
- b. I am probably about average on detail orientation.
- c. Not really; I tend to overlook small details or fail to do a really thorough job of attending to the details required on many tasks or jobs.
- d. Definitely no; I often miss important details on a task or job and I'm much better at things requiring little or no detail work.

31. How organized are you when working on a job or task?

- a. I plan and organize my work almost to a fault.
- b. I am quite well organized on most tasks and jobs.
- c. Though I generally get the job done, I must admit my organization could usually be better.
- d. Frankly, I am very poor at planning and organizing, and prefer to play it by ear when working on a task or job.

32. While you were growing up, which of the following types of persons did you especially admire? (Circle all that apply.)

- a. Religious leaders.
- b. Counselors or social workers.
- c. Coaches.
- d. Military leaders.
- e. Political leaders.
- f. Parents.
- g. Businessmen.
- h. Doctors.
- i. Lawyers.
- j. Union leaders.
- k. Movie or TV stars.
- l. Teachers.
- m. Tough guys.
- n. None of the above.

33. When you were in high school, did you participate in any of the following clubs, societies, or activities? (Circle all that apply.)

- a. Dramatics, debating, or speech clubs.
- b. Fraternity or social groups.
- c. Music, band, chorus, orchestra, etc.
- d. History or foreign language clubs.
- e. Math or science clubs.

- f. Literary, magazine, or newspaper.
 - g. Team sports or other sports.
 - h. Student government.
 - i. None of the above.
34. In which of the following groups or social organizations did you hold office before coming on recruiting duty (president, secretary, chairman of the committee, etc.)? (Circle all that apply.)
- a. Athletic and recreational clubs--bowling, golf, tennis, chess, bridge, photography, etc.
 - b. Fraternal and cultural societies--Elks, Masons, K of C, IOOF, YMCA, college fraternity, dramatics, debating, bible class, etc.
 - c. Civic and political organizations--Rotary, Kiwanis, Lions, Chamber of Commerce, Young Republicans, American Legion, PTA, etc.
 - d. Student government.
 - e. Never held any offices.
35. Which of the following were you active in during high school? (Circle all that apply.)
- a. National Honor Society.
 - b. Future Farmers of America.
 - c. Boy Scouts or Girl Scouts.
 - d. Church youth group.
 - e. Future Teachers of America.
 - f. Cheerleaders group.
 - g. Marching band.
 - h. Football team.
 - i. Basketball team.
 - j. Baseball team.
 - k. Golf team.
 - l. Tennis team.
 - m. Hockey team.
 - n. Wrestling or boxing.
 - o. Gymnastics team.
 - p. Swimming team.
 - q. Public service organizations or volunteer work (such as in hospitals).
 - r. None of the above.
36. Listed below is a series of 24 statements about different things you may or may not have done and feelings you may have had about various situations and people. Please read each one carefully and decide whether or not the statement is true or mostly true about you. Circle any statement that is more true about you than it is false.
- a. You graduated from high school.
 - b. You have been suspended or expelled from school.
 - c. You dropped out of high school.
 - d. You worked instead of going to school.
 - e. As a child, you hated being ordered around at home or at school.
 - f. You ran away from home once.
 - g. As a child, you hated working around the house.

- h. Your parents may have tried, but they did a poor job.
- i. It bothers you to have your family move around as much as military families do.
- j. You worked full time before you were sixteen.
- k. You had a social security card when you were sixteen.
- l. Obeying strict rules and regulations bothers you.
- m. It bothers you to be taken off jobs before they are finished.
- n. You try to do well in things that don't interest you.
- o. You were never fired from a job.
- p. You have a bank account.
- q. From twelve to fourteen, you usually went around with a group of three or four guys.
- r. You went to a beer party before you were sixteen.
- s. You smoked regularly before you were sixteen.
- t. You have hitchhiked more than 200 miles.
- u. You have been in a gang fight.
- v. You have been arrested.
- w. You were in jail overnight once.
- x. It is important to you to be stationed near a large city.

REVISION OF THE BIOGRAPHICAL SURVEY
(Example)

29. Some people easily become involved in a task while others seldom really "dig into" a task or job. How involved do you usually become in a task or job?

- a. I often have trouble sticking with it; other things almost always seem to come up to distract my attention.
- b. I sometimes become involved in a task or job that interests me greatly, but most of the time I quickly lose interest.
- c. I often become heavily involved in a task or job provided it's of interest to me.
- d. I almost always become engrossed in tasks or jobs.

Forty-three individuals in the pretest sample responded (c) to this question, while only three responded (a), five responded (b) and six responded (d). Since the item was clearly not differentiating well among persons in the sample, two of the alternatives were changed in an attempt to increase variability in responding to the item. The revised item appears below.

29. Some people easily become involved in a task while others seldom really "dig into" a task or job. How involved do you usually become in a task or job?

- a. I often have trouble sticking with it; other things almost always seem to come up to distract my attention.
- b. I sometimes become involved in a task or job that interests me, but if I'm not very interested, I seldom become involved.
- c. I often become heavily involved in a task or job.
- d. I almost always become engrossed in tasks or jobs.

APPENDIX E

DESCRIPTION OF THE MONTE CARLO CROSS-VALIDATION PROCEDURE

DESCRIPTION OF THE MONTE CARLO CROSS-VALIDATION PROCEDURE

The Monte Carlo cross-validation procedure has been developed over a period of several years by Dr. Rodney L. Rosse of the research staff of Personnel Decisions Research Institute. Its purpose is to provide cross-validity estimates that are stable in comparison to other sample-splitting procedures commonly used to index the cross-validity of sets of predictors in test validation research.

This appendix first reviews briefly the critical issues related to making cross-validity estimates, discusses the approach taken by the Monte Carlo procedures to address these issues, and outlines how these procedures were applied to data from the study described in this report.

Most methods for estimating the cross-validity of predictors involve "sample-splitting" in one form or another. The basic procedure used in these "sample-splitting" techniques is characterized as follows:

1. The total sample is divided into two subsamples at random; one subsample may be called a "developmental" subsample and the other may be called a "cross-validation" subsample.
2. The set of statistical rules applied to the total sample to obtain the prediction function is fully repeated on the "developmental" subsample, resulting in an "interim" prediction function that is only to be used for cross-validity estimation.
3. This "interim" function is applied to the computation of point predictions for each subject in the "cross-validation" subsample.
4. The appropriate measure of association (usually a Pearsonian coefficient) is computed; that is considered the estimate of validity (i.e., cross-validity) for the prediction function.

Clearly, the validity estimated in this manner is conditional upon the occurrence of a particular split and thus is neither a unique nor an efficient estimator. It is not unique because simple repetition of the process with a new random split will result in a different cross-validity estimate. Therefore, there is "within-sample" variability in the statistic. It is not efficient because the within-sample variability contributes to the random sampling distribution of the validity estimate.

The Monte Carlo cross-validation procedure provides a unique and efficient estimator of the validity of an empirically determined prediction function by using the computational capability of a scientific computer to do the above four steps repeatedly with rerandomization of the "developmental" and "cross-validation" split for each repetition, and reestimation of the constants for the prediction function. The process gives rise to a distribution of "cross-validities," the median or mean of which may be taken as the estimates of validity (referred to here as the Monte Carlo cross-validity estimator).

The number of repetitions is determined by the standard error of the mean of the "within-sample" distribution. According to the central limit theorem, this standard error decreases as a function of the reciprocal of the square root

of the number of cross-validity estimates obtained, i.e., the number of repetitions. Therefore, the researcher may simply specify the maximum within-sample error he wishes to have in the average of the estimators of validity, and then repeat the cross-validation procedure a sufficient number of times to reduce the standard error below that level. Thus, within the limits of this error, the final validity estimate is a unique and efficient variance estimator.

Now to comment briefly on an approach to selecting predictor measures in test validation research, an approach that influences significantly the outcome of Monte Carlo analyses. It has proven generally useful to select a relatively few predictors with maximum a priori potential for predicting performance. In particular, the following "rules" have been used in conducting validation research with the Monte Carlo method:

1. The variables to be introduced into the estimation process are kept to a minimum and restricted to those for which there is some a priori reason to believe that they would be useful (even if just a "hunch").
2. The manner in which the variable or item is to be used in the prediction function is predetermined as much as possible (e.g., a variable may be constrained to weight only in a positive or negative direction or a variable may be assumed to have value only if included with a "curvilinear" fit).
3. The weighting that is used may be specified arbitrarily by the researcher instead of depending upon the empirical estimation of weights (e.g., arbitrary unit or differential weighting of items or scales).
4. Items that are to be empirically keyed may be rescaled and several response categories lumped together when indicated by judgment on an a priori basis.
5. The use of multiple regression and, particularly, step-wise multiple regression, is avoided where data are insufficient to obtain highly stable estimates of the covariances among predictor variates. This is because these regression methods are greatly influenced by the covariance estimates and would not be expected to replicate where inadequate data are used in the estimation.

In the analysis of data for the present project, there was a particularly good basis for a priori keying of items and scales because the information obtained in the pilot phase was available to aid the researchers in making these keying decisions. Therefore, two keying strategies, one for each of the two different types of predictor data, were used to make a priori predictor-criterion matches. The types of data were (1) biographical history items, and (2) interest inventory items.

The biographical history items were empirically keyed (for each criterion) according to the following rules:

1. A response (or group of responses) for a particular item was weighted +1, -1, or 0 according to whether the response was hypothesized to be positively or negatively related (or unrelated) to performance on the criterion.

2. To be included in a Monte Carlo run, the mean of the criterion scores for those endorsing a response was required to fall outside of the 48 to 52 range in T-score values (T-score means = 50, standard deviations = 10) and the mean criterion score was required to be in the proper direction.

3. And finally, in order for a response to be considered for inclusion in a Monte Carlo run, the proportion of respondents who endorsed it was required to be greater than or equal to .05 of the half sample.

For each Monte Carlo run, then, item responses were selected on the basis of these rules and the resultant key was cross-validated on the other half of the sample. The cross-validity coefficient for each run was formed by computing a single predictor score for each individual in the cross-validation sample using the key developed on that run and correlating those scores with the performance scores obtained by the same individuals. The median of these cross-validity coefficients then formed our final validity estimate for a key that was developed using total sample data.

In the procedure used for estimating cross-validity for the interest inventory, each item was treated as an ordered scale with three points, corresponding to the like-indifferent-dislike format of the Strong Campbell Interest Inventory used in the study. A priori judgments about items were made according to pretest data analysis results and conceptual considerations. For these a priori selections of items for each criterion, it was specified that each item should be considered positively related, negatively related, to have no predetermined direction but to be included in the keys, or not to be included in the keys. Thus, for each criterion a pool of items was developed that appeared to show promise for predicting performance on that criterion. These a priori judgments were based upon an analysis of pretest data and conceptual considerations. Further, most but not all of the items selected in this manner were given directional constraints.

For each Monte Carlo run, a Pearsonian correlation coefficient between each item in the pool and the criterion was computed using half the sample. Items were included in the key by the following steps:

1. The direction of the correlation for an item was required to be consistent with the a priori constraints.

2. The items were rank ordered according to the absolute values of correlations with the criterion, and the top 50 percent were selected for the key.

As with the biographical items, for each Monte Carlo run, prediction scores were computed for members of the cross-validation sample using the key developed for that run and these scores were correlated with the criterion performance scores. Again, the median of these cross-validity coefficients served as the final validity estimate for a key formed on the basis of total sample data.

APPENDIX F

HYPOTHESIZED PREDICTOR-PERFORMANCE CRITERION RELATIONSHIPS

HYPOTHESIZED PREDICTOR-PERFORMANCE CRITERION RELATIONSHIPS

The following tables present hypothesized relationships between the inventory scales and items selected for predictor composites and the pretest performance criteria. In all of the tables, the intended direction of the item is provided: a minus sign (-) indicates a negative direction; a plus sign (+), a positive direction; and the numeral one (1), a keyed item with no direction given. For the Monte Carlo program, the (+) and (-) keyed items were constrained to be keyed in the intended direction, while (1) items could be keyed in either direction in the Monte Carlo runs. See Appendix E for a description of the Monte Carlo method. In the Marine Corps sample, the keys presented here were cross-validated directly. Thus, that key does not include the items keyed (1).

Table F-1

Hypothesized PRF, CPI, DPQ, and SDI Item-Performance Criterion Relationships

	Prospecting & Selling	Planning & Organizing	Good Relations with Community	Overall Performance	Productivity
Personality Research Form (PRF)	+		+	+	
Exhibition Order		+			+
Dominance	+	+		+	+
Sociability			+	+	
Social Presence			+		
Socialization			+		
Achievement via Conformance		+	+		
Good Impression				+	
Social Closeness			+		
Hard Work		+	+		+
Authoritarianism		+			
Impulsiveness		-			
Sales Effectiveness	+			+	+

^aThe pretest performance Prospecting and Selling corresponds to the main sample criterion Salesmanship Skills, Planning and Organizing to Administrative Skills, and Good Relations with the Community to Human Relations Skills.

Table P-1 (Continued)

	Prospecting & Selling ^a	Planning & Organizing	Good Relations with Community	Overall Performance	Productivity
Intelligence					
Supervisory Qualities					
Initiative					
Self-Assurance					
Perceived Oc- cupational Level					
Decision- making Ability					
Sociometric Popularity					
Perceived Maturity					
Self Descrip- tion Inventory (SDI)					+

^aThe pretest performance Prospecting and Selling corresponds to the main sample criterion Salesmanship Skills; Planning and Organizing, to Administrative Skills; and Good Relations with the Community, to Human Relations Skills.

Table F-2

Hypothesized SCII Item-Performance Criterion Relationships

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
1			-
2			
3	+		
4	+	+	+
5	+	+	+
6	+	+	
7	+	+	
8			
9			
10			
11			
12			
13			
14		+	
15		+	
16			
17	+		
18	+		
19			
20			
21		+	
22			
23	1		1
24	1	+	1
25	+	+	
26	1		
27			
28			
29			

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
30	+	1	1
31	1	1	1
32	-	1	1
33	1	1	1
34			+
35			
36	+	+	
37	+	+	
38			
39	+	+	
40		1	
41		1	
42	-	+	1
43	1	1	-
44	+	+	
45	1	1	1
46			
47			
48	-		
49	+		
50			+
51	-	-	
52			
53	1	+	1
54	+	1	1
55			-
56	+		
57		+	
58		-	
59			
60			
61			

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
62		+	
63	+		
64	1	+	
65		-	
66			-
67			
68	1		-
69			-
70	1		
71			
72	+	+	
73	1		1
74			
75			
76		+	
77			
78			-
79	1		
80		-	
81		+	
82		-	-
83	-	-	-
84			
85			-
86	+	+	
87			
88			
89	-	-	
90			
91			
92		+	-
93	+	+	+

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
94		+	
95			-
96		+	
97		-	
98			
99	+		
100	-		-
101		+	
102			
103			
104			
105	1		+
106		-	
107	1		
108			
109			
110			
111	1		
112	-		
113	1	+	+
114			
115			
116	-		
117			
118	-	-	-
119			
120			
121	1	+	+
122			
123	1	-	1
124	+		

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
125			
126			
127			
128	+	+	+
129	-		
130			
131			
132	+		
133	-		-
134	-	-	-
135			
136	+	+	
137			
138	1		
139	1		
140	-		
141			
142	-		-
143			
144	1		
145			
146	-		
147	1	+	
148	1		-
149	-		
150			
151	-		
152			
153			-
154			
155		+	
156			

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
157			
158			-
159	-		-
160			
161	-		
162			
163	-		-
164			-
165	-	-	
166			
167			
168	1	-	-
169	-		
170			
171			
172			
173			
174	1	-	-
175		-	-
176			
177		+	-
178			-
179	+	+	+
180			
181	-		
182			
183	-		-
184		-	
185		+	
186	-	-	
187			
188	+	+	

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
189	+	+	+
190	-		-
191			-
192			+
193			
194	-		
195			
196			
197	-		-
198	-	-	-
199			
200	+		
201		-	-
202			
203		-	
204			
205	-		-
206			
207			
208	-		
209	1		
210			-
211	+	+	
212			+
213			
214			-
215			
216			
217			-
218		-	
219		-	
220			

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
221		+	
222	+		
223			-
224			
225			
226			
227	-		-
228			-
229			-
230			
231			
232			-
233			-
234	1	+	
235			
236			
237			
238			-
239	1		
240			-
241			
242	1		
243		-	-
244	1		-
245			-
246	+		
247		+	
248			
249	-		-
250		+	
251	+	+	
252	+	+	

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
253			
254			-
255	+		
256			-
257			
258		+	
259			
260			-
261			
262			-
263			
264			
265	-	-	-
266			-
267	1		-
268			-
269	1		
270			-
271			
272			
273			
274	1		
275	-	-	-
276	-		-
277	-		-
278	-		-
279			
280	-		
281			
282			-
283	+	+	
284			

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
285	-	-	-
286			
287			
288	+		
289			+
290	+	+	
291			
292			
293			
294			
295			
296			+
297			-
298	-		
299			
300			
301	+		
302	1		
303	+	+	
304			
305			
306			
307	1		
308	-		
309	-	-	
310			
311	+		
312	-	-	-
313	-		
314	1		-
315			

Table F-2 (Continued)

Item	Prospecting & Selling	Planning & Organizing	Good Relations with Community
316			+
317		+	
318			-
319	-		
320	-	-	
321			
322			
323	1		1
324	-	-	-
325			

Table F-3

Hypothesized Biographical Survey Item-Performance Criterion Relationships

Item	Prospecting & Selling & Two Overall Perform- ance Criteria		Good Relations with Community		Planning & Organizing	
	^a +	-	+	-	+	-
1						
2						
3						
4						
Section 1						
1						
2						
3						
4						
5						
6						
7						
8			3 ^b	1, 2		
9						
10					5	1, 2
11				5		
12		5				5
13						
14						
15						
16				4		
17	1				1	
18	1	2	1	2	1	2
19	1					
20						
21						

^a(+) means the response was keyed +, (-) it was keyed -.

^bRefers to item response alternative.

Table F-3 (Continued)

Item	Prospecting & Selling & Two Overall Perform- ance Criteria		Good Relations with Community		Planning & Organizing	
	^a +	-	+	-	+	-
22			1	2		
23	2	1	2	1	2	1
24		3	1			
25	1	3, 4		3, 4	1	3, 4
26						
27		3	2			
28	1	4, 5	1	4, 5	1	4, 5
29						
30						
31						
32	4, 5		4, 5		4, 5	
33	2		2		2	
34	1	5	1	5	1	5
35	17		17		17	

^a(+) means the response was keyed +, (-) it was keyed -.

APPENDIX G

IMPLEMENTATION OF THE PREDICTOR TEST BATTERY
FOR USE IN SELECTING NAVY AND MARINE CORPS RECRUITERS

**IMPLEMENTATION OF THE PREDICTOR TEST BATTERY
FOR USE IN SELECTING NAVY AND MARINE CORPS RECRUITERS**

Final composites of personality scales have been developed for each criterion. Separate composites were developed for the Navy and Marine Corps using data from the total samples of Navy and Marine Corps recruiters. That is, the best and most complete information available about relationships between the personality scales in the battery and performance on the various criteria was used to select final sets of scales for each criterion for the Navy and another series of final sets of scales for each criterion for the Marine Corps. This selection of personality scales stands the greatest chance of providing valid predictor composites for each of the performance criteria because it makes use of the maximum amount of empirical information available about personality scale-performance relationships.

Likewise, the SCII and Biographical and Opinion Survey items that appeared most promising for predicting recruiter effectiveness on each performance criterion were selected according to relationships between item responses and performance on each criterion. Separate selections were made for the Navy and Marine Corps using data from the total samples. Thus, the personality composites and keys for the SCII and Biographical and Opinions Survey are ready to go. These composites and keys are not the same ones as those listed in Appendix F, the hypothesized predictor-criterion relationships. They have been drawn from that list and contain only the items and scales that the study found to be empirically related to recruiter performance.¹ This means that the predictor battery may be shortened from its present length for operational use, since several of the items and scales did not prove to be valid and therefore may be eliminated from the battery. The reduction in length should cut testing time to approximately 1-1/2 hours.

How can the Navy or Marine Corps use the predictor battery in practice? The battery should be used both for diagnostic purposes and for helping to make final selection decisions on candidates for the recruiter job. First, scoring the battery should prove straightforward. Templates that fit over answer sheets can easily be constructed, allowing for quick scoring of the battery. Two clerks, working independently of one another, should score the responses for each candidate and check each other's scores to ensure the accuracy of final scores. Further, the raw scores should be transformed to Navy standard scores (T-scores with a mean of 50 and standard deviation of 10) to make the predicted performance levels more interpretable for Navy and Marine Corps decision makers. The clerks could also make this transformation using a simple conversion table.

Alternatively, the battery could be machine-scored at a central location. This approach would require that respondents use machine-scorable answer sheets when completing the battery.

Either scoring approach is capable of yielding standardized predicted performance scores for candidates on each of the three performance dimensions--

¹The final composites and keys are not displayed in this report for security reasons. They have been delivered to NAVPERSRANDCEN, Navy Recruiting Command, and officials at Marine Corps Headquarters.

Selling, Human Relations, and Administrative Skills. For the Marine Corps, such scores can also be computed for the criteria Overall Performance and Production, while a score on the criterion Overall Performance (but not Production) can be estimated for the Navy. These predicted performance scores can be plotted on a display sheet for each candidate in a manner similar to that shown in the example in Figure G-1. This information can then be used both diagnostically and as an aid in making a final select or reject decision.²

To point up its potential use as a diagnostic tool, suppose that a Navy recruiting district had a couple of locations in which the Navy was not well accepted, and it was deemed important to get into those locations recruiters who possessed good human relations skills. If the district saw fit to place new recruiters into those locations, district decision makers might be well advised to attend to candidates' predicted performance in the human relations area. The candidate whose scores are depicted in Figure G-1, for example, would not be a very good choice for such a placement. Of course, the general level of the profile and predicted effectiveness in the Overall Performance criterion should obviously be weighted heavily in selecting Navy or Marine Corps recruiters.

²The Marine Corps should not include predicted performance scores for both Overall Performance and Production on a display sheet because the information may be confusing. Since these two criteria are both summary measures of effectiveness, the predicted scores for the two criteria should be averaged and displayed as a single score.

Name A. J. Candidate

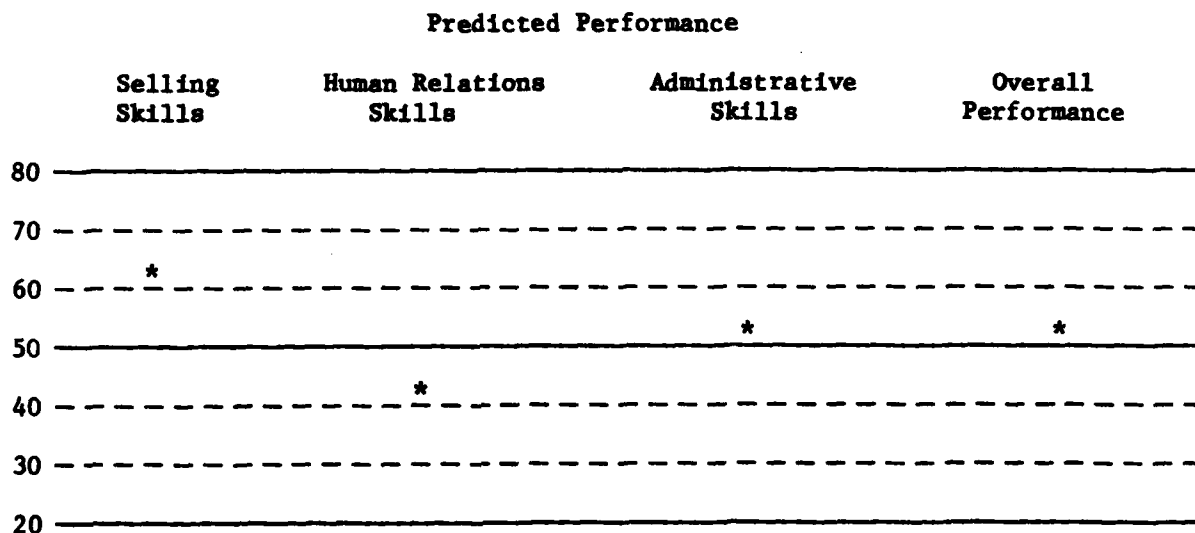


Figure G-1. Example of a possible design for a display system to depict a candidate's predicted performance as a recruiter.

DISTRIBUTION LIST

Chief of Naval Operations (OP-102) (2), (OP-11), (OP-987H)
Chief of Naval Material (NMAT 08T244)
Chief of Naval Research (Code 450) (4), (Code 458) (2)
Chief of Information (OI-2252)
Director of Navy Laboratories
Commandant of the Marine Corps (Code MPI-20)
Chief of Naval Education and Training (N-5)
Commander, Naval Military Personnel Command (NMPC-013C)
Commander, Navy Recruiting Command (Code 00), (Code 11), (Code 20), (Code 22),
(Code 50)
Commander, Navy Recruiting Area One
Commander, Navy Recruiting Area Three
Commander, Navy Recruiting Area Four
Commander, Navy Recruiting Area Five
Commander, Navy Recruiting Area Seven
Commander, Navy Recruiting Area Eight
Technical Library, Air Force Human Resources Laboratory (AFSC),
Brooks Air Force Base
Army Research Institute for the Behavioral and Social Sciences (Reference
Service)
Military Assistant for Training and Personnel Technology, Office of the
Under Secretary of Defense for Research and Engineering
Defense Documentation Center (12)