POST READING & WRITING TEST VALIDATION RESEARCH

1994



THE COMMISSION ON PEACE OFFICER STANDARDS AND TRAINING

STATE OF CALIFORNIA

POST READING & WRITING TEST

VALIDATION RESEARCH

John Weiner Sr. Personnel Selection Consultant

November 1994

CALIFORNIA COMMISSION ON PEACE OFFICER STANDARDS AND TRAINING

THE MISSION OF CALIFORNIA COMMISSION ON PEACE OFFICER STANDARDS AND TRAINING IS TO CONTINUALLY ENHANCE THE PROFESSIONALISM OF CALIFORNIA LAW ENFORCEMENT IN SERVING ITS COMMUNITIES.

Copyright California Commission on Peace Officer Standards and Training

Published November 1994

*Revised July 1998

*This revision reflects only Commissioner's page and Preface page with updated Executive Directors signature.

For information about copies of this publication, contact:

POST Media Distribution Desk 1601 Alhambra Boulevard Sacramento, CA 95816 (916) 227-4856

COMMISSIONERS

Collene Campbell, Chairman

Public Member

Rick TerBorch Vice Chairman

David C. Anderson

Sherman Block

Charles S. Brobeck

Michael T. Carre

Philip del Campo, Ph.D.

Ted Hunt

Thomas J. Knutson, Ph.D.

William B. Kolender

Ronald E. Lowenberg

Daniel E. Lungren

Jan Scully

7.4

Barbara G. Warden

Kenneth J. O'Brien

Chief Arroyo Grande Police Department

Sergeant Los Angeles County Sheriff's Department

Sheriff Los Angeles County

Chief Irvine Police Department

Investigator Orange County District Attorney's Office

Public Member

Director Los Angeles Police Protective League

Professor of Communication Studies California State University, Sacramento

Sheriff San Diego County

Chief Huntington Beach Police Department

Attorney General

District Attorney Sacramento County

Council Member, 5th District City of San Diego

Executive Director

i



PREFACE

Reading and writing abilities are important for the successful performance of patrol officer work. POST has conducted several studies spanning more than a decade which have documented the central role of written expression and comprehension in the performance of patrol officer duties in California, and which have established the empirical validity of POST's tests of these abilities, the Entry-Level Law Enforcement Officer Reading & Writing Test Battery. This report fully describes a recent empirical validation study of the POST Reading & Writing Tests and summarizes several previous POST studies.

The research results provide important substantiation of the validity of POST's Reading & Writing Tests with reference to the prediction of several criterion measures of job performance -- measures which extend throughout the initial phases of an entry-level officer's career, including performance in the basic academy, field training, probation, and beyond probation as a tenured officer. Also important is the finding that the tests are *not* unfair to various racial/ethnic and gender subgroups in predicting job performance. Additional insights are provided regarding the validity of alternative configurations of the tests along with an Essay Test of writing ability, the negligible effects of several potential moderators upon test validity, and the utility of the test battery -- in most instances, significant gains in employee performance may be expected as the Reading & Writing Test cut score is increased.

In general, our findings are consistent with cumulative research evidence pertaining to law enforcement occupations and are offered in support of POST's reading and writing selection standards for entry-level peace officers.

D'Brien

KÈNNETH J. O'BRIEN Executive Director

a

ACKNOWLEDGEMENTS

The present study was completed with the assistance of five agencies: the Los Angeles, Oakland, Sacramento, San Diego, and San Francisco Police Departments. POST extends its thanks to these agencies for providing staff to complete academy and job performance evaluations for the hundreds of officers in the study. The following individuals were especially helpful in coordinating project activities at their respective agencies: Sergeant John Kelley, San Francisco PD; Sergeant Chris Ellis and Officer Diane Wendell, San Diego PD; Sergeants Jim Richart and Jim McNair, Los Angeles PD; Lieutenants Max Canfield and Richard Shiraishi, Sacramento PD; and Sergeant Patrick Haw, Oakland PD.

The study was conducted under the general direction of Dr. John Berner, Chief of the POST Standards and Evaluation Services Bureau, who provided technical guidance throughout the project, as well as in producing the final report. Richard Honey, Sr. Personnel Selection Consultant, developed the report writing criterion rating scales and coordinated the collection of academy instructor ratings of students' report writing. William Dyer, Personnel Selection Consultant, was involved in the development of the research data base and provided reliability estimates for the Reading & Writing Test forms. Finally, Vickie Pruden, Office Technician, provided assistance in producing the criterion rating instruments, coordinating data collection efforts with local agencies, and in typing the final report.

TABLE OF CONTENTS

COMMISSION ON PEACE OFFICER STANDARDS AND TRAINING	i
PREFACE i	ii
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS v	ii
LIST OF TABLES	ci
INTRODUCTION	1
METHOD	3
Research Design	3 3
Predictors	4 4 2
Criteria 1 Academy Performance Measures 1 Field Training Performance Measures 1 Data Collection 1	2 3 5 8
ANALYSIS AND RESULTS 2	1
Sample Characteristics	1
Prediction of Academy Instructor Ratings 2 Descriptive Statistics 2 Validity Evidence 2	4 4 4
Prediction of Overall Academy Success/Failure 2 Descriptive Statistics 2 Validity Evidence 2	8 8 8
Prediction of Academy Proficiency Test Scores 3 Descriptive Statistics 3 Validity Evidence 3	2 2 2

Prediction of Field Training Officer (FTO) Ratings	35
Descriptive Statistics	35
Validity Evidence	39
Prediction of Field Training Success/Failure	41
Descriptive Statistics	41
Validity Evidence	44
	an a' sa
Differential Prediction Analysis	46
Academy Instructor Ratings	46
Academy Success/Failure	49
Academy Proficiency Test Scores	52
FTA Ratings	52
Summory of Validity Evidence	56
Oberall Desults	56
	50
Within-Group Validity	50
Differential Prediction	J. 7 ,
	60
Analysis of Alternative Test Batteries	00
Individual Test Validities	60
Relative Difficulty of Individual Tests	61
Multiple Regression Analyses	63
Alternative Test Batteries	64
Validity Evidence for Alternative Batteries	64
Relative Difficulty of Alternative Batteries	65
Summary	69
Moderator Analysis	69
Time Span Effect	69
Potential Rating Errors	71
Rater Characteristics	73
Summary	77
Life lity	77
Utility in Predicting Academy Instructor Ratings	78
Utility in Predicting Academy Success/Failure	82
Utility in Prodicting Academy Praticiency Test Scores	84
Halikte in Dradiating FTA Patings	84
Othery III Freedomy F EO Ramigs	88
Summary,	00
Citi Margaretti Oukon Dangaretti	005
Comparison of Hindings with Uther Research	00
Other PUS I Studies	20
Cumulative Research	91

viii

	SUMMARY AND CONCLUSIONS 99
	Reading & Writing Test Scores Predict Academy and Job Performance 99
	Overall Results
	Differential Prediction
v.	Alternative Test Batteries Offer Little or No Gain in Prediction
2	Moderator Variables Found to Have No Effect Upon Validity
	Reading & Writing Test Scores Offer Utility
	Conclusions
	REFERENCES
	APPENDICES
	Appendix A: Scoring Procedure for the Reading & Writing Test Battery
	Appendix B:
	Essay Test Instructions and Scoring Procedure
	Appendix C:
· · ·	Basic Academy Report Writing Instructor Rating Booklet - End of Writing Instruction Evaluation
	Appendix D
	Basic Academy Success/Failure Data Collection Form and Coding Instructions
	Appendix E:
	FTO Rating Booklet: Final Evaluation
a 7	Appendix F: Field Training Success/Failure Data Collection Instructions and Coding Sheet
	Annendix G
ан Са. — — — — — — — — — — — — — — — — — — —	Multiple Regression Analyses: Reading & Writing Test and Essay Test Predicting
	Basic Training and Field Training Performance
• •	Appendix H:
en de la composition de la composition de la composition de la Composition de la composition de la comp	Empirical Expectancy Tables: Reading & Writing Test Scores Predicting Academy Graduation vs. Failure/Withdrawal for Academic Reasons - 1987 study
	ix.
	f , f ,

/

LIST OF TABLES

		~
Table 1	Descriptive Statistics for Reading & Writing Tests	6
Table 2	Reading & Writing Test Intercorrelations	7
Table 3	Internal Consistency Reliability Estimates for Reading & Writing Tests	8
Table 4	Test-Retest Reliability Estimates for Reading & Writing Tests 1	.0
Table 5	Descriptive Statistics for the Essay Test 1	.1
Table 6	Basic Academy Success/Failure Outcomes	4
Table 7	Field Training Success/Failure Outcomes	.7
Table 8	Breakdown of Subjects for Whom Test Scores and Performance Data were	
	Collected by Agency, Race/Ethnicity, and Gender	2
Table 9	Breakdown of Subjects for Whom Reading & Writing Test and Proficiency	
	Test Scores were Collected by Academy 2	23
Table 10	Descriptive Statistics - Reading & Writing Test and Essay Test Scores	
	(Academy Instructor Rating Sample) 2	25
Table 11	Descriptive Statistics - Academy Instructor Ratings	25
Table 12	Academy Instructor Rating Intercorrelations	26
Table 13	Zero-Order Correlations - Reading & Writing Test and Essay Test Scores	
	with Academy Instructor Ratings 2	27
Table 14	Descriptive Statistics - Reading & Writing Test and Essay Test Scores	
	(Academy Success/Failure Sample)	29
Table 15	Frequency Distribution of Basic Academy Outcomes	60
Table 16	Zero-Order Correlations - Reading & Writing Test and Essay Test Scores	
	with Academy Success/Failure	51
Table 17	Descriptive Statistics - Reading & Writing Test and Essay Test Scores	
	(Academy Proficiency Test Sample)	3
Table 18	Descriptive Statistics - Academy Proficiency Test Scores	33
Table 19	Zero-Order Correlations - Reading & Writing Test and Essay Test Scores	-
	with Academy Proficiency Test Scores	34
Table 20	Descriptive Statistics - Reading & Writing Test and Essay Test Scores	-
20010 20	(FTO Rating Sample)	6
Table 21	Descriptive Statistics for FTO Ratings	16
Table 22	FTO Rating Intercorrelations	37
Table 23	Characteristics of the FTO (Rater) Sample	18
Table 24	Zero-Order Correlations - Reading & Writing Test and Essay Test Scores	Ū
	with FTO Ratings	IU
Table 25	Descriptive Statistics - Reading & Writing Test and Essay Test Scores	.0
	(Field Training Success/Failure Sample)	12
Table 26	Frequency Distribution of Field Training Outcomes	12
Table 20	Zero Order Correlations – Reading & Writing Test and Essay Test Sector	15
Table 27	with Field Training Success/Feilure	15
Table 20	With Field Hamming Success/Famule	IJ
rable 28	Differential Prediction Analysis by Race/Ethnicity and Gender - 10tal	17
	kead/write Score Predicting Academy Instructor Kating	¥/-

Table 29	Differential Prediction Analysis by Race/Ethnicity and Gender - Essay Test Score Predicting Academy Instructor Rating	48
Table 30	Differential Prediction Analysis by Race/Ethnicity and Gender - Total Read/Write Score Predicting Academy Success/Failure	50
Table 31	Differential Prediction Analysis by Race/Ethnicity and Gender - Essay Test	51
Table 32	Differential Prediction Analysis by Race/Ethnicity and Gender - Total	51
Table 33	Differential Prediction Analysis by Race/Ethnicity and Gender - Essay Test	23
т. Т. С. С. С. С. С.	Score Predicting Academy Proficiency Test Score	54
Table 34	Differential Prediction Analysis by Race/Ethnicity and Gender - Total	
a the state of the	Read/Write Score Predicting FTO Ratings	55
Table 35	Differential Prediction Analysis by Race/Ethnicity and Gender - Essay Test	. e
a de la compañía de l	Score Predicting FTO Ratings	57
Table 36	Summary of Validity Evidence for Reading and Writing Tests	58
Table 37	Relative Difficulty of Reading & Writing Tests - Mean Z-scores by	
	Race/Ethnicity and Gender	62
Table 38	Validity Evidence for Alternative Reading and Writing Test Batteries	66
Table 39	Within-Group Validity Evidence for Alternative Reading and Writing Test	
1	Batteries	67
Table 40	Relative Difficulty of Alternative Reading & Writing Test Batteries - Mean	
	Z-scores by Race/Ethnicity and Gender	68
Table 41	Moderator Analysis - Time Between Testing and Criterion Data Collection .	70
Table 42	Moderator Analysis - Potential Sources of Rater Bias in FTO Ratings	72
Table 43	Semipartial Correlations - Read/Write Total Scores and Adjusted FTO	
•	Ratings of Writing Ability	72
Table 44	FTO Mean Ratings of Writing Ability and Overall Performance by	
	Evaluator Race/Ethnicity and Gender	74
Table 45	FTO Mean Ratings of Writing Ability and Overall Performance by	
	Evaluator and Ratee Race/Ethnicity	75
Table 46	FTO Mean Ratings of Writing Ability and Overall Performance by	
	Evaluator and Ratee Gender	76
Table 47	Empirical Expectancy Tables - Reading & Writing Test Scores Predicting	
	Academy Instructor Ratings - Total Sample	79
Table 48	Empirical Expectancy Tables - Reading & Writing Test Scores Predicting	
i e e e	Academy Instructor Ratings by Race/Ethnicity and Gender	81
Table 49	Empirical Expectancy Table - Reading & Writing Test Scores Predicting	
	Academy Success/Failure - Total Sample	83
Table 50	Empirical Expectancy Tables - Reading & Writing Test Scores Predicting	
	Academy Proficiency Test Scores - Total Sample	85
Table 51	Empirical Expectancy Tables - Reading & Writing Test Scores Predicting	
	Proficiency Test Scores by Race/Ethnicity and Gender	86
Table 52	Empirical Expectancy Tables - Reading & Writing Test Scores Predicting	
a.	FTO Ratings - Total Sample	87

xii

Table 53	Empirical Expectancy Tables - Reading & Writing Test Scores Predicting	
·	FTO Ratings by Gender	89
Table 54	Summary of Validity Evidence for POST Reading & Writing Tests	92
Table 55	Summary of Within-Group Validity for the POST Reading & Writing	
	Test Battery	94
Table 56	Summary of Differential Prediction Analyses of the POST Reading &	
	Writing Test Battery	96

14 A.

INTRODUCTION

The POST Entry-Level Law Enforcement Officer Examination is a battery of five tests designed to measure fundamental reading and writing abilities. The tests are offered by POST to local agencies as a means of complying with statewide selection standards for California's entry-level peace officers [cf. POST reg. 1002(a)(9)].

The initial form of the Reading & Writing Test Battery was developed over a dozen years ago (Honey & Kohls, 1981) and has since been the subject of continuous research (Honey, 1983; Weiner & Berner, 1987). In 1990, POST initiated a joint research study to examine the empirical validity of both the Entry-Level Reading & Writing Test Battery and the POST Work Sample Test Battery (a battery of physical abilities tests). This report describes the methodology and results of that research pertaining to the Reading & Writing Test Battery,¹ along with a ten-year retrospective study of Reading & Writing Test score predictions of students' performance in the basic academy. An overview of previous POST Reading & Writing Test validation research is also presented, along with a brief description of cumulative job analytic and empirical validity evidence for verbal ability tests in predicting performance in law enforcement occupations.

¹A report describing the research pertaining to the physical abilities tests was published separately by POST (Weiner, 1993).

METHOD

Research Design

The study followed a predictive criterion-related strategy wherein the POST Reading & Writing Test Battery was administered to job applicants or newly hired basic academy students, and, after some time, measures of the examinees' performance in the basic academy and subsequent field training were obtained. The predictive nature of the tests was examined by computing Pearson product-moment correlation coefficients between test scores and the measures of academy and field training performance.

The research included a ten-year retrospective study of the relationship between Reading & Writing Test scores and basic academy students' subsequent scores on a standardized achievement test measuring their knowledge of the basic course curriculum; i.e. the POST Basic Academy Proficiency Test. Further details regarding the achievement test are given below (see section entitled "Criteria").

In addition to examining overall test validity, analyses were conducted to assess the fairness of the battery in predicting job performance for racial/ethnic and gender subgroups, the validity and relative difficulty of alternative test batteries, the validity-moderating effects of several variables, and the practical utility of Reading & Writing Test scores.

Sample

Five large police departments agreed to participate in the portion of the research that entailed collecting specially developed measures of academy and field training performance. These included the Los Angeles, Oakland, Sacramento, San Francisco, and San Diego police departments. Each of these departments operates its own basic academy. Officers were selected as subjects for the study if they attended one of these basic academies ending between June 1990 and November, 1991.²

Subjects of the retrospective study of Academy Proficiency Test performance were selected as a result of retrieving their scores from POST's computer files that are maintained as part of the operational Reading & Writing Test and Academy Proficiency Test programs. Test scores were available for a ten-year period, spanning 1983 to 1992. The research sample. represented a total of 35 agencies, including the above five agencies and 30 basic training academies. Characteristics of the research sample are described later in the Analysis and Results section.

²Officers from Los Angeles PD were further sampled by selecting all female academy students and a random sample of male students as needed to achieve an overall target of 150 officers. The emphasis on selecting females was for purposes of conducting the physical abilities test research.

Predictors

General Test Characteristics

Reading & Writing Test Battery. The POST Reading & Writing Test Battery is comprised of five subtests, including two reading comprehension tests and three tests designed to measure facets of writing ability. The tests are objectively scored and all but one employ a traditional multiple-choice format; the last test (Cloze Test) utilizes a specially designed format designed to capture examinees' free responses to fill-in-the-blank questions.

The tests include:

1.

- **Clarity** (15-items): This test is designed to measure the ability to write sentences that are unambiguous and free from errors that might obscure the meaning of written communication. The examinee is presented a pair of sentences and instructed to pick the one that is more clearly written.
- 2. **Spelling** (15 items): This test measures the ability to correctly spell common words. The examinee is presented a sentence with a word deleted, followed by a list of several alternative spellings of the deleted word. The examinee is to identify the correct spelling of the word from among the listed alternatives.
- 3. Vocabulary (15 items): This test is designed to measure the ability to understand and appropriately use common words. The examinee is presented a sentence with one word underlined, followed by several listed words. The examinee is to identify from among the alternatives the one that most nearly matches the meaning of the underlined word.
- 4. **Reading Comprehension** (20 to 30 items):³ This test is designed to measure the ability to derive meaning from written English. The test employs a format that has traditionally been used in personnel selection: the examinee is presented a brief passage followed by several questions regarding facts contained in the passage and interpretation of the information. The passages contained in the test were designed to be representative of the types and level of reading materials commonly encountered on the job.⁴
- 5. Cloze Test (40 items): This test is designed to measure reading ability in a manner that reflects the cognitive processes involved in reading. The examinee is presented a passage in which words have been systematically deleted and

⁴See 1981 POST study.

³Forms 200 and 210 contain 30 multiple-choice Reading Comprehension items, while forms 220 and higher contain 20 such items.

replaced by a blank line. The examinee is to determine the missing word based on his or her knowledge of the English language in conjunction with the context provided by the total passage. The rationale for cloze testing is further described in the initial test validation report (Honey & Kohls, 1981).

The tests are administered with a single overall time limit of 2-1/2 hours. Several alternate forms of the battery are in existence, and forms 200, 210, 220, 230 and 240 were used in the present study. Scores on the test are reported on a T scale, which is calibrated to an original research sample obtained in 1983. Appendix A contains the specific scoring procedure for the battery.

General descriptive statistics for Reading & Writing Test forms 200-240 are reported in Tables 1 and 2, including means, standard deviations (SDs), and intercorrelations. Values for subsequent forms (which are not included in the present study) are also shown in Table 1. These statistics are based upon scores obtained by over 120,000 examinees tested between 1983 and 1993. The vast majority of these scores were obtained by job applicants in an employment selection context; a small percentage (6%) were obtained by academy students or prospective students.⁵

Estimates of the internal consistency reliability (coefficient alpha) of Reading & Writing subtest and composite scores are presented in Table 3. This index of reliability represents the degree of precision with which examinees' "true" abilities are measured by their scores on the given sample of items. Values of the reliability coefficient may range from 0 (no reliability) to 1.0 (perfect reliability). Reliabilities for forms 200-240 and additional forms are shown which are based on a sample of approximately 60,000 job applicants and academy · students/prospective students tested between 1989 and 1993.⁶ It should be noted that the Cloze Test was *not* included in these reliability estimates since these items are not independent (a necessary condition for this type of reliability index). Therefore, the composite test reliabilities shown in the table, while in the acceptable range of .80s to .90s, should be viewed as *underestimates* of the reliability of the total battery.

⁵These data were retrieved from POST's computer data files which are maintained as part of the statewide testing program. Examinees' first-obtained test scores were selected (it is not uncommon for individuals to take the test more than once). A small percentage of test scores were identified as "outliers" and excluded from the analysis, as follows: if Read/Write total <10 or >76; if average percent score on writing tests=0, or average percent score on reading test=0, or cloze test score=0. Application of these rules resulted in the exclusion of examinees representing all racial/ethnic and gender groups.

⁶It was necessary to select a smaller sample of examinees due to the excessive computer resource requirements associated with the computation of internal consistency estimates of reliability (i.e., examinees' responses to each item must be analyzed). The sample was selected from the above sample of over 120,000 examinees, dating back to 1989.

			Tab	le 1		ч ^к .	
Descrip	tive	Statistics	for	Reading	&	Writing	Tests

Prot Parme	I matel	Dead	T wraite				1.00	
1est Form	R/W	STD STD	STD	Clarity	Spen	Vocad	M/C Read	Cloze
Form 200 (N=23,761)	1							
Mean	48.50	50.14	47.08	70.81	73.75	76.43	69.39	65.11
SD	12.35	12.31	11.81	15.23	17.03	15.24	17.67	14.17
Form 210 (N=32,846)	1							
Mean	45.80	48.35	44.07	69.90	73.44	68.55	67.65	62.88
SD	12.43	12.63	11.90	14.09	16.51	16.00	17.79	14.64
Form 220 (N=26,633)								
Mean	47.14	50.47	44.34	70.08	73.09	69.52	72.17	63.33
SD ·	12.23	12.48	11.41	13.98	15.99	15.51	18.30	14.02
Form 230 (N=24,517)			a substantia					
Mean	48.60	50.47	46.94	72.92	77.88	69.78	70.72	64.77
SD	13.22	12.90	12.51	14.91	16.64	18.26	16.24	17.09
Form 240 (N=2,700)					· · · · ·			i i i i i i i i i i i i i i i i i i i
Mean	48.64	50.61	46.87	67.08	79.41	73.86	69.99	65.83
SD	12.91	11.67	13.18	15.77	17.81	19.31	14.72	15.82
Form 250 (N=11,542)								
Mean	45.34	48.80	42.78	69.07	72.08	66.81	68.68	63.17
SD	11.88	11.41	11.84	15.01	16.75	17.59	15.47	14.55
Form 260 (N=1,736)		· · ·				- - -		
Mean	47.35	53.51	41.67	67.18	69.21	68.42	74.94	67.68
SD	11.33	11.32	11.18	13.80	17.72	14.27	14.94	14.38
Form 270 (N=814)							· · · · · · · · · · · · · · · · · · ·	
Mean	49.78	54.07	45.44	72.03	70.94	73.07	73.30	70.62
\$D	11.98	11.97	11.31	13.09	16.90	14.93	16.63	14.79
Form 400 (N=1,504)	1	· · · · · · · · · · · · · · · · · · ·						
Mean	45.46	48.09	43.71	68.34	68.99	73.47	67.72	62.20
SD	12.40	12.18	11.90	15.44	17.37	15.29	17.33	13.78
Form 440 (N=209)	a an							
Mean	45.76	49.60	42.75	67.11	72.31	68.45	69.47	63.98
SD	13.33	11.72	14.01	15.37	20.10	20.16	14.65	16.34

Note: Scores are based on examinees between 1983 and 1993; approximately 94% were job applicants and 6% were academy students or prospective students. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores.

	Total R/W	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze
Total Read/Write								
Reading STD	.92							
Writing STD	.91	.66						
Clarity	.66	.48	.74	-				
Spelling	.63	.40	.76	.35				
Vocabulary	.76	.63	.77	.38	.35			
M/C Reading	.83	.92	.60	.43	.36	.56	••	
Cloze	.82	.89	.60	.43	.37	.57	.63	

Table 2Reading & Writing Test Intercorrelations

Note: N=126,262. All correlations are significant (p<0.001). Scores are based on job applicants and academy students tested on forms 200-270, 400 and 440 between 1983 and 1993. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores.

Test Form		Sub	otests		Composites		
	Clarity	Vocab	Spelling	M/C Reading	Writing ^a	M/C Read & Writing ^b	
Form 200 N=3,735	.575	.744	.764	.875	.847	.919	
Form 210 N=6,469	.474	.598	.588	.821	.739	.866	
Form 220 N=10,604	.465	.587	.602	.714	.739	.820	
Form 230 N=20,208	.504	.668	.676	.680	.775	.839	
Form 240 N=2,700	.525	.763	.782	.671	.832	.870	
Form 250 N=11,542	.458	.674	.625	.616	.736	.800	
Form 260 N=1,736	.411	.513	.661	.650	.711	.789	
Form 270 N=814	.448	.651	.654	.738	.774	.851	
Form 400 N=1,504	.565	.648	.637	.805	.766	.871	
Form 440 N=209	.491	.741	.766	.635	.818	.854	

Table 3Internal Consistency Reliability Estimates for
Reading & Writing Tests

Note: Based on job applicants and academy students tested between 1989 and 1993. Examinees' first-obtained scores were selected; outliers were excluded. Internal consistency estimates were not computed for Cloze Test since item responses are not independent.

⁸Coefficient alpha for 45-item score based on multiple-choice items contained in the three writing subtests (Clarity, Vocabulary, Spelling).

^bReliability estimate for linear composite of Writing score plus multiple-choice Reading score [see Guilford, (1954) <u>Psychometric</u> <u>Methods</u>, 2nd edition. McGraw-Hill, p.393]. Table 4 contains test-retest reliability estimates for the five Reading & Writing subtests and the total test battery. This index of reliability represents the extent to which an obtained test score provides a stable estimate of an examinee's ability over time (where ability is defined relative to other examinees). These reliability estimates are based upon scores obtained by job applicants and academy students who took the same test form twice between 1983 and 1993. The retest reliabilities for operational Read/Write total scores were found to be in the mid-.80s, indicating that such scores provide reasonably stable estimates of ability.

Essay Test. Scores on an Essay Test of writing ability were also obtained for purposes of this study. The Essay Test is not part of POST's operational Reading & Writing Test Program, but was included for purposes of examining alternative tests.

The Essay Test entails instructing examinees to "Describe an event that made a significant change in your life. Explain why that event had importance for you." Examinees are then given 40 minutes to write an essay addressing this topic. The essays are scored using a 6-point holistic scale, where each point on the scale is anchored by a description of an essay exemplifying a particular level of writing ability (with ability increasing monotonically from 1 to 6). The scoring process is normative in that essays are scored relative to the selected essays. Each essay is scored independently by two raters; borderline competent essays are scored by a third rater (i.e., cases where one rater scored the paper as a "4," which denotes basic competence, and a second rated scored the paper as a "3," which denotes less than competent writing); papers that receive two ratings that differ by two or more points are also rated by a third rater. A sample Essay Test and scoring guide are contained in Appendix B.

General descriptive statistics for the Essay Test are presented in Table 5, including distributions of original ratings and final scores, and the mean and SD of final scores. These statistics are also based upon an applicant sample which includes individuals who were not part of the present validation study.⁷ The interrater reliability for this sample of Essay scores is estimated to be .933.⁸ It is noteworthy that 38% of the applicants received final scores below the level of "Basic Competence" (sum of two ratings less than 8). Reliability estimates for Essay scores obtained by the validation analysis samples described later in this report were lower (approximately .840), likely due to restriction in the range of scores (the subjects were pre-selected on the basis of existing employee selection procedures).

⁷Essay Test scores were retrieved from POST research files for a sample of 818 applicants at Sacramento City Police Department, obtained in 1987. The essays were scored by POST staff (Richard Honey and Luella Luke) and Sacramento PD staff (Capt. Steve Segura).

⁸Spearman-Brown estimate applied to correlation between two independent ratings. In those instances where 3 ratings were made, the highest and lowest values were selected. Note that this estimate does not take into account the operational scoring process which sometimes involves a third rater.

Table 4Test-Retest Reliability Estimates for
Reading & Writing Tests

Test Form	Clarity	Vocab	Spelling	M/C Read	Cloze	Total R/W
Form 200 N=2029	.670	.721	.730	.718	.715	.869
Form 210 N=2618	.602	.670	.635	.754	.691	.843
Form 220 N=2440	.627	.724	.706	.724	.688	.871
Form 230 N=1890	.580	.757	.710	.632	.751	.858
Form 240 N= 114	.657	.822	.767	.607	.838	.849
Form 250 N= 644	.600	.809	.692	.629	.745	.858

Note: Based on job applicants and academy students who took the same test form twice between 1983 and 1993. Examinees' first two obtained scores on the same form were selected. Outliers were excluded (total score <10 or >76, or segment score=0, or mean score for entire administration <30). Cases were also excluded if the time between test and retest was zero or greater than 2000 days, or if the data were suspected to be duplicate (i.e., 3 or more examinees on the same test dates each obtained the same score on both test administrations).

Table 5Descriptive Statistics for the Essay Test

Rating Level	Ratir	ng #1	Rati	ng #2	Rating #3	
	Freq	Pct	Freq	Pct	Freq	Pct
1	33	4.0%	15	1.8%		
2	120	14.7%	66	8.1%	2	1.9%
3	216	26.4%	189	23.1%	40	37.7%
4	328	40.1%	311	38.0%	56	52.8%
5	99	12.1%	180	22.0%	8	7.5%
6	22	2.7%	57	7.0%		
Total	818		818		106	

Original Ratings

Final Score

Score	Freq	Pct
2	15	1.8%
3	18	2.2%
4	50	6.1%
5	65	7.9%
6	163	19.9%
8	278	34.0%
9	99	12.1%
10	80	9.8%
11	28	3.4%
12	22	2.7%
Total	818	

Mean Score	SD
7.43	2.15

Note: Essay Test scores retrieved from POST research files for sample of 818 applicants at Sacramento City Police Department, obtained in 1987. Rating #1 was the lowest and #2 the highest obtained rating per applicant. A third rating was to be made as needed to resolve a discrepancy of 2 or more points between ratings, or between ratings of "3" and "4" (basic competence).

Data Collection

Of the five agencies that agreed to provide academy and field training performance data, two (Sacramento and San Francisco) had previously administered the Reading & Writing Test Battery to job applicants, and these scores were collected for the present study. The Read/Write Test was administered experimentally at the remaining three agencies (Los Angeles, Oakland, and San Diego) to their basic academy students. The Essay Test was previously administered to job applicants at one of the five agencies (Sacramento) and the scores were collected for use in the study. The Essay Test was administered on an experimental basis to academy students at the remaining four agencies.⁹

As indicated earlier, Read/Write Test scores were collected for the retrospective study from computer files that are maintained in conjunction with the POST Reading & Writing Test program. Read/Write scores were retrieved for all examinees for whom Academy Proficiency Test scores were also available (specific procedures for selecting test records are described later). In those instances where an individual had taken the Reading & Writing Test more than once, the first test administration was selected. Hence, some of the test dates are several years prior to the dates that students attended basic training. The test dates ranged from August 1983 to March 1992.

Scoring. The Reading & Writing Tests were scanned and scored by POST staff using equipment and software from the operational testing program. The Essay Tests administered in Sacramento were scored by an independent consultant (Dr. Charles Moore), and the remaining essays were scored by assistants of Dr. Moore, as well as POST staff.

Criteria

A total of five criterion measures were collected for purposes of examining the predictive validity of Reading & Writing Test scores. Three of the criterion measures reflected performance in *basic training*, including: (1) academy instructor ratings of students' demonstrated writing abilities, (2) students' overall success or failure in completing the basic academy, and (3) students' scores on the POST Basic Academy Proficiency Test, a paper-and-pencil achievement test. The remaining two criterion measures were designed to reflect performance on the job, in *field training*, including: (4) Field Training Officer (FTO) ratings of officers' job performance and demonstrated writing abilities, and (5) officers' overall success or failure in completing field training. Each of these performance measures, with the exception of the POST Proficiency Test, was developed specifically for purposes of this study. The performance measures are briefly described below.

⁹The Essay Test was administered in conjunction with the Reading & Writing Test Battery at the Los Angeles, Oakland and San Diego academies.

Academy Performance Measures

Academy Instructor Ratings. A rating instrument was developed to assess students' writing ability following completion of the report writing curriculum in basic training. Four salient components of writing ability were assessed which were identified by POST staff and academy writing instructors:

- 1. **Organization and Narrative**: The ability to write clear and organized narrative in reports.
- 2. Writing Mechanics: The ability to write reports that are free of errors in fundamental writing mechanics (i.e., reports that are characterized by good grammar, punctuation, spelling and word choice).
- 3. **Information and Elements:** The ability to include all necessary information and elements in reports.
- 4. **Timeliness**: The ability to write acceptable reports in a timely manner.

A 5-point rating scale was developed for each writing component. The scale points were labeled 1=Very Poor, 2=Poor, 3=Adequate, 4=Good, and 5=Excellent, and each point was anchored by a description of writing corresponding to that level. Appendix C contains a sample Academy Instructor Rating Booklet.

Academy Success/Failure. Information regarding students' success or failure in completing basic training was collected from their respective academies using a specially developed coding sheet. Success or failure in completing training, including reasons for failure, were each represented by a special code that was recorded for each student in the study. In general, the interest was in identifying students who were terminated or resigned while performing unsatisfactorily for reasons related to reading and writing abilities (e.g., inadequate report writing skills, analytical skills, job knowledge, or other skills and abilities). However, in order to obtain a more complete picture of students' academy performance, other reasons for failure were recorded (e.g., inadequate physical ability or weaponless defense skills, attitude, work habits, etc.).

The academy success/failure outcomes and their corresponding codes are shown in Table 6. A sample data collection form and coding instructions are contained in Appendix D.

Table 6 Basic Academy Success/Failure Outcomes

Completed Training:

C1=Graduated in normal time.

C2=Graduated academy but required extra time (remediation).

Resigned:

R1=Overall academy performance was satisfactory.

Unsatisfactory performance due to . . .

R2=inadequate report writing skills.

R3=generally inadequate analytical skills.

R4=inadequate weaponless defense skills.

R5=inadequate baton skills.

R6=failure to complete POST physical conditioning program or Work Sample Test Battery.

R7=inadequate knowledge, skills and abilities <u>other than</u> report writing, analytical and physical (e.g., knowledge of laws, procedures, tactics; learning ability; oral communication skills; driving skills; etc.).
 R8=other reasons (e.g., attitude, motivation, work habits).

Terminated:

T1=Overall academy performance was satisfactory.

Unsatisfactory performance due to . . .

T2=inadequate report writing skills.

T3=generally inadequate analytical skills.

T4=inadequate weaponless defense skills.

T5=inadequate baton skills.

T6=failure to complete POST physical conditioning program or Work Sample Test Battery.

T7=inadequate knowledge, skills and abilities other than report writing, analytical and physical (e.g.,

knowledge of laws, procedures, tactics; learning ability; oral communication skills; driving skills; etc.). T8=other reasons (e.g., attitude, motivation, work habits).

Injury: Il=Cadet withdrew because of an injury.

Recycled: Cadet was recycled to attend next academy due to

N1=injury or illness.

N2=inadequate report writing skills.

N3=generally inadequate analytical skills.

N4=inadequate weaponless defense skills.

N5=inadequate baton skills.

N6=failure to complete POST physical conditioning program or Work Sample Test Battery.

N7=inadequate knowledge, skills and abilities other than report writing, analytical and physical (e.g.,

14

knowledge of laws, procedures, tactics; learning ability; oral communication skills; driving skills; etc.). N8=other reasons (e.g., family emergency).

Academy Proficiency Test. All students attending a POST-certified basic training academy are required to complete the POST Basic Academy Proficiency Test, a written multiple-choice test designed to measure knowledge of the basic course curriculum. The test is criterion-referenced, in that each item is designed to measure a specific performance objective in the basic course curriculum. Several forms of the test have been developed and forms 4 - 8 were included in the present study. Total scores on the Proficiency Test are equated using a linear transformation procedure and then scaled to a T-score distribution (mean=50, SD=10).

Criterion-related validity evidence was obtained for the Proficiency Test in a previous POST study (Weiner & Berner, 1987), wherein academy students' test scores were found to be significantly predictive of their subsequent job performance as measured by specially developed and administered job simulations, performance ratings, and indices of successful completion of field training and probation. Further details regarding this research, as well as the relationship of the Proficiency Test to the POST Basic Course curriculum, are provided in the technical report for the above-referenced 1987 study..

Field Training Performance Measures

FTO Ratings. A rating instrument was developed to obtain assessments of each officer's performance in field training, to be completed by the officer's designated field training officer (FTO). The instrument contained the same four writing ability scales used in the above described academy rating instrument, along with Likert-type scales designed to elicit ratings of performance quality for 19 physical job tasks and abilities, 18 knowledges, skills, abilities and traits/characteristics (KSATs), and a global dimension representing overall job performance. The performance ratings were intended to reflect the officers' performance throughout field training and resulting final level of performance.¹⁰

The physical performance rating scales were adapted from a previous POST study (Weiner, 1988) and were based upon statewide job analyses conducted by POST (Kohls, Berner & Luke, 1979; Berner & Kohls, 1982; Berner, et al., 1985). The KSAT scales were also adapted from a previous POST study (Weiner & Berner, 1987) and are based upon dimensions of work identified in the 1979 job analysis.¹¹ Appendix E contains a sample FTO Rating Booklet.

The following knowledge and ability scales were selected for the analysis, along with the four writing ability scales, as potentially relevant to prediction by measures of reading and writing ability:

¹⁰A short form of the FTO rating instrument was also developed which did not include scales for the 18 job elements for purposes of the Physical Abilities Test component of the research data collection effort. This form was used to collect periodic ratings throughout the field training program due to the relatively infrequent nature of some physical job tasks.

¹¹The knowledge scales are reflective of major content domains contained in the Basic Course.

Legal Knowledge: Demonstrate working knowledge of laws, codes, and legal procedures (e.g., accurately detect crimes and violations and apply all appropriate codes; comply with legal requirements when making arrests, conducting searches, and obtaining evidence; write reports that include all necessary legal elements).

Knowledge of Departmental Policies and Procedures: Demonstrate working knowledge of department policies, regulations and procedures (e.g., able to verbalize and apply them appropriately).

Knowledge of Patrol Procedures: Demonstrate working knowledge of procedures and techniques for performing patrol activities (e.g., able to verbalize and apply appropriate methods for beat patrol, suspect approach, vehicle stop, searching, restraining, prisoner transportation, and handling different types of calls).

Knowledge of Investigative Procedures: Demonstrate working knowledge of procedures and techniques for gathering information (e.g., able to verbalize and apply appropriate methods for locating and identifying victims, witnesses, and suspects; interviewing; collecting and preserving evidence).

Learning: Comprehend new information and apply that which has been learned on the job.

Overall Field Training Success/Failure. A data collection form was developed to describe officers' overall success or failure in completing field training in a manner similar to the above described academy success/failure coding instrument. A different coding scheme was used to describe field training outcomes and there were some differences in the types of outcomes to be recorded. The field training outcomes to be recorded and their respective codes are listed in Table 7. A sample data collection form and coding instructions are contained in Appendix F.

Table 7 Field Training Success/Failure Outcomes

Completed:

C1=Completed in normal time. C2=Completed -- required extra time (remediation). C3=Completed -- time required unknown.

Resigned (voluntary):

R1=Overall job performance was satisfactory.

Unsatisfactory performance due to . . .

R2=inadequate physical ability. R3=inadequate report writing skills. R4=inadequate analytical skills.

R5=inadequate job knowledge, skills or abilities <u>other than</u> physical/report writing/analytical. R6=other reasons (e.g., attitude, motivation, work habits). R7=performance level unknown.

Terminated (involuntary):

T1=Overall job performance was satisfactory.

Unsatisfactory performance due to . . .

T2=inadequate physical ability. T3=inadequate report writing skills. T4=inadequate analytical skills. T5=inadequate job knowledge, skills or abilities <u>other than</u> physical/report writing/analytical. T6=other reasons (e.g., attitude, motivation, work habits). T7=performance level unknown.

Failed But Continued in Non-Patrol Assignment:

Unsatisfactory performance due to . . .

F2=inadequate physical ability.

F3=inadequate report writing skills.

F4=inadequate analytical skills.

F5=inadequate job knowledge, skills or abilities <u>other than</u> physical/report writing/analytical. F6=other reasons (e.g., attitude, motivation, work habits).

Other:

O1=Injury. O2=Other (retired, transferred, etc.).

Data Collection

Academy Performance Data. Academy Instructor Rating Booklets and Academy Success/Failure Coding Sheets were distributed to designated project coordinators at each academy prior to the ending dates of the academy classes (between June 1990 and November 1991). POST staff personally reviewed the data collection procedures with each agency coordinator. Coordinators were instructed to obtain two independent instructor ratings of each student's writing ability whenever possible. The completed ratings and coding sheets were returned to POST upon which time POST staff reviewed the forms for accuracy and completeness. Follow-up calls were made to the agencies as needed to clarify or to obtain additional data.

Proficiency Test scores obtained by basic academy students between January 1984 and February 1992 were retrieved from computer data files maintained as part of the POST Proficiency Test Program. These students' Read/Write Test scores were then retrieved from POST's Reading & Writing Test Program computer files. Scores on the two tests were matched and selected according to specific criteria.¹² Scores were retrieved for a total of 13,347 examinees, including Reading & Writing forms 200-240 and Proficiency Test forms 4-8. Essay Test scores obtained in the present study were matched to Proficiency Test scores for 227 examinees.

Field Training Performance Data. FTO Rating Booklets and Field Training Success/Failure Coding Sheets were distributed to local agency project coordinators shortly after the graduation date of each academy class. POST staff personally reviewed all data collection procedures with each agency coordinator. Prior to the administration of the FTO rating booklets, POST staff provided on-site training to FTOs in which they were instructed in the rating procedure and completed a rating calibration exercise. A final evaluation of each trainee's performance throughout field training was made by his or her last assigned FTO at the end of the field training program.¹³

The completed field training performance ratings and overall success/failure coding sheets were mailed to POST, and staff reviewed the data forms to ensure that they were completed properly. Again, follow-up correspondence with agencies was made as needed to clarify or to obtain additional data. The final ratings were completed between August 1990 and November 1992.

¹²Proficiency Test scores obtained between Jan84 and Feb92 were merged with Read/Write scores obtained between Aug83 and Aug91, with the following restrictions: (a) excluded cases with more than one Proficiency Test score, (b) selected first obtained Read/Write score, (c) excluded cases if time between Read/Write Test and Proficiency Test date was less than 4 months or greater than 36 months; (d) excluded Proficiency Test scores if less than 17.5; (e) excluded Read/Write scores if less than 10 or greater than 76; (f) excluded data from miskeyed (re-scanned) administrations.

¹³The field training programs of the participating agencies varied in length, as follows: San Diego, 10 weeks; Oakland, 15 weeks; San Francisco, 12 weeks; Sacramento, 24 weeks; and LAPD, 6 to 12 months (their OJT phase is not formally defined as "field training"). Field trainees were also rated by their respective FTOs at 4- or 5-week intervals in order to obtain information regarding less frequently observed physical job activities (only final FTO ratings were obtained for LAPD trainees).

Data Entry and Verification. All performance data were key-entered by POST Information Services Bureau staff. Sample data records were printed and independently verified to ensure the accuracy of the computer data files.
ANALYSIS AND RESULTS

This section describes an extensive series of analyses addressing the validity of test scores and related issues. Following a brief description of characteristics of the research sample, the results of a series of validity analyses are presented corresponding to the various criterion measures of performance that were obtained; namely, academy instructor ratings, overall academy success/failure, Academy Proficiency Test scores, FTO ratings, and overall field training success/failure. Analyses of racial/ethnic and gender group test score predictions of academy and field training performance are then described, as evidenced via a statistical procedure referred to in this report as "Differential Prediction Analysis."

Several alternative test batteries were constructed consisting of different combinations of the Reading & Writing Tests and the Essay Test. Results are described with regard to overall and within-group validity, as well as the relative difficulty of the alternative batteries for racial/ethnic minorities and females.

Next, analyses are described regarding the moderating effects of several variables that were identified as having the potential to affect the obtained validity results; i.e., time between predictor and criterion data collection, indices of potential rater bias, and characteristics of the rater.

The practical utility of Read/Write Test scores is then described in terms of empirical expectancy tables depicting relative gains in academy and field training performance associated with test scores. Overall and within-group results are presented.

Finally, the results of the present research are compared to the findings of previous POST validation studies, as well as cumulative job analytic and empirical validity evidence for verbal ability tests in predicting performance in law enforcement occupations.

Highlights of the findings of the present study are presented in the next chapter (see "Summary and Conclusions").

Sample Characteristics

Breakdowns of the subjects for whom test scores, academy performance, and job performance data were collected are presented in Table 8, by agency, race/ethnicity, and gender. The majority of performance ratings and success/failure data were collected from Los Angeles, San Diego and San Francisco. While the majority of the research subjects were male (68% to 86%) and White (59% to 72%), representation of racial/ethnic minorities (American Indians, Asians, Blacks, Filipinos, and Hispanics) and females was sufficient to facilitate some subgroup analyses. As seen in Table 9, Proficiency Test data were obtained for students from many academies in addition to the five agencies that provided performance ratings.

21

Table 8Breakdown of Subjects for WhomTest Scores and Performance Data were Collectedby Agency, Race/Ethnicity, and Gender

			Academy	Performance	,		I	Field Trainin	g Performa	ince
	Inst R: Freq	ructor atings Pct	O Succe Freq	verall sss/Failure Pct	Profi T Freq	ciency est Pct	FTO Freq	Ratings Pct	O Succe Freq	verall ss/Failure Pct
Agency										
Oakland	54	10.7%	49	11.6%	297	2.2%	15	4.7%	45	10.7%
Los Angeles	155	30.8%	101	23.9%	1,007	7.5%	87	27.0%	136	.32.2%
Sacramento	42	8.3%	29	6.9%	203	1.5%	7	2.2%	17	4.0%
San Diego	141	28.0%	138	32.6%	493	3.7%	118	36.6%	121	28.7%
San Francisco	112	22.2%	106	25.1%	235	1.8%	95	29.5%	102	24.2%
Other					11,112	83.3%				
Race/Ethnicity ^b										
Am. Indian	1	0.2%		· · · · · · · · ·	137	1.0%				
Asian	28	5.6%	27	6.4%	407	3.1%	21	6.4%	24	5.7%
Black	68	13.6%	58	13.8%	996	7.6%	47	14.3%	68	16.2%
Filipino	7	1.4%	7	1.7%	154	1.2%	3	0.9%	8	1.9%
Hispanic	86	17.2%	64	15.2%	1,821	13.8%	49	14.9%	66	15.7%
White	303	60.5%	258	61.3%	9,495	72.2%	205	62.3%	248	59.0%
Other	8	1.6%	· · · 7	1.7%	148	1.1%	3	0.9%	6	1.4%
Gender	1									
Male	356	71.1%	342	81.2%	11,149	85.6%	236	71.7%	285	68.0%
Female	145	28.9%	79	18.8%	1,882	14.4%	93	28.3%	134	32.0%
Total	504		423	ang tanàn ang	13,347		329		422	

^aAgency not reported for 7 FTO ratings and 1 field training pass/fail record.

^bRace/ethnicity not reported for 3 academy instructor ratings, 2 academy pass/fail records, 1 FTO rating, 2 field training pass/fail records, and 189 Proficiency Test records.

^cGender not reported for 3 academy instructor ratings, 2 academy pass/fail records, 3 field training pass/fail records, and 316 Proficiency Test records.

Table 9Breakdown of Subjects for WhomReading & Writing Test and Proficiency Test Scores were Collected
by Academy

Academy	Freq	Pct
Alameda County Sheriff's Academy	360	2.7
Allan Hancock College	118	0.9
Butte Center	562	4.2
California Department of Forestry	2	0.0
California Department of Parks and Recreation	8	0.1
California Highway Patrol	1196	9.0
Central Coast Academy	261	2.0
Fullerton College Basic Academy	13	0.1
Golden West College	734	5.5
Kern County Academy	64	0.5
Long Beach Police Department	141	1.1
Los Angeles County Sheriff's Department	1021	7.6
Los Angeles Police Department	1007	7.5
Los Medanos College	615	4.6
Modesto Regional Training Center	548	4.1
Napa Valley College	192	1.4
Oakland Police Department	297	2.2
Orange County Sheriff's Department	483	3.6
Redwoods Center	174	1.3
Rio Hondo College	346	2.6
Riverside City College	512	3.8
Sacramento Criminal Justice Training Center	116	0.9
Sacramento County Sheriff's Academy	483	3.6
Sacramento Police Department Academy	203	1.5
San Bernardino County Sheriff's Academy	423	3.2
San Diego County Sheriff's Academy	251	1.9
San Diego Police Department	493	3.7
San Francisco Police Department	235	1.8
San Joaquin Delta College	241	1.8
San Jose Community College District	618	4.6
Santa Rosa Center	504	3.8
Southwestern College, San Diego	151	1.1
State Center, Fresno	424	3.2
Tulare-Kings County Academy	252	1.9
Ventura County Police and Sheriff's Academy	299	2.2
Total	13,347	

23

Prediction of Academy Instructor Ratings

Descriptive Statistics

Predictors. Means and standard deviations (SDs) for Reading & Writing Test and Essay scores obtained by the academy rating analysis sample are shown in Table 10.

Criterion. Sample means and SDs on the Academy Instructor Rating Scales are shown in Table 11. Intercorrelations between independent ratings by two different instructors on each of the writing component scales are shown in Table 12.¹⁴ The estimated interrater reliabilities of each scale ranged from .62 (Timeliness) to .75 (Organization/Narrative and Writing Mechanics). The estimated reliability of the mean instructor rating was .77.¹⁵

The mean elapsed time between Reading & Writing testing and completion of the academy ratings was approximately 16 months.¹⁶

Validity Evidence

All Reading & Writing Test scores (total and subtest) and Essay Test scores were found to be significantly predictive (p<.0001) of mean instructor ratings of demonstrated writing ability, as well as of each of the four writing ability scales. The individual test correlations with mean writing ability rating ranged from .24 (Clarity) to .33 (multiple-choice Reading Comprehension). Correlations with the individual ability scales ranged from .19 (Clarity with Timeliness) to .32 (multiple-choice Reading with Organization/Narrative). The obtained zero-order correlations for Reading & Writing Test and Essay Test scores with academy instructor ratings are shown in Table 13.

¹⁴A Principal Components analysis of these data was conducted indicating that a single factor accounts for 74% of the total variance in instructor ratings; all scales loaded highly (.83 to .88) on the first factor. Thus, it is reasonable to combine the ratings into a single composite score.

¹⁵Spearman-Brown formula applied to interrater correlations.

¹⁶Average difference between academy ending date and Reading & Writing Test date. Subjects were excluded from the analysis if they were not tested prior to the end of basic training.

	111	Table 10	
		Descriptive Statistics	
Readi	ing &	Writing Test and Essay	Test Scores
	(Acad	emy Instructor Rating S	ample)

	Total R/W	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay
Mean	51.97	54.21	49.21	74.78	77.74	74.94	76.85	67.40	8.19
SD	10.68	10.66	10.52	13.55	14.32	14.17	15.70	12.63	1.63

Note: Read/Write Test N=504; Essay N=413. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores; Essay score is a raw sum of two ratings.

Table 11Descriptive StatisticsAcademy Instructor Ratings

	Mean Writing Ability	(1) Org & Narrative	(2) Writing Mechanics	(3) Info. & Elements	(4) Timeliness
Mean	3.68	3.69	3.74	3.76	3.53
SD	0.63	0.73	0.74	0.70	0.74

Note: N=504. Writing ability scale: 5=Excellent, 4=Good, 3=Adequate, 2=Poor, 1=Very Poor.

25

D / //1	Rater #2							
Kater #1	1. Org & Narrative	2. Mech.	3. Info & Elements	4. Time	Mean Writing (1-4)			
1. Org & Narrative	.60 (.75)							
2. Mechanics	.43	.60 (.75)						
3. Info & Elements	.46	.45	.56 (.72)					
4. Timeliness	.36	.42	.38	.45 (.62)				
Mean Writing Ability (1-4)	.84	.86	.86	.85	.63 (.77)			

 Table 12

 Academy Instructor Rating Intercorrelations

Note: N=404. Interrater reliability estimates shown on diagonal in parentheses, derived via Spearman-Brown formula. All correlations are significant (p<.0001).

	Academy Rating:								
Test:	Mean Rating	1. Org & Narrative	2. Mech	3. Info & Elements	4. Time				
R/W Total	.39	.35	.32	.37	.32				
Read STD	.36	.34	.31	.33	.28				
Write STD	.35	.29	.28	.34	.29				
Clarity	.24	.20	.20	.25	.19				
Spelling	.28	.25	.23	.26	.25				
Vocabulary	.26	.21	.20	.27	.23				
M/C Reading	.33	.32	.28	27	.26				
Cloze	.31	.27	.25	.31	.23				
Essay	.31	.31	.23	.26	.26				

Table 13Zero-Order CorrelationsReading & Writing Test and Essay Test Scores with
Academy Instructor Ratings

Note: Read/Write Test N=504; Essay Test N=413. All correlations are significant (p<.0001, one-tailed).

Prediction of Overall Academy Success/Failure

Descriptive Statistics

Predictors. Means and SDs of Reading & Writing Test and Essay Test scores for the academy success/failure analysis sample are shown in Table 14.

Criterion. A frequency distribution of academy pass/fail outcomes is contained in Table 15. Outcome data were obtained for a total of 467 students, of whom 89.1% successfully completed basic training. Very few subjects in the study were identified as having resigned while performing unsatisfactorily or terminated for reasons related to inadequate report writing skills; only 4 subjects (0.5%) were identified as such. Moreover, no students were identified as having failed to complete basic training for reasons related to inadequate analytical skills. Thus, there was little variation in student success/failure related to reading and writing abilities and as a result, little to be predicted by Reading & Writing test scores (the obtained correlations would be expected to be small under such conditions).

An overall index of academy success/failure was constructed to serve as a criterion measure of performance for the validity analyses. The index was designed to reflect overall successful performance versus failure to complete training for reasons that would be expected to be germane to a test of reading and writing abilities, including inadequate report writing, job knowledge and other abilities (e.g., knowledge of laws, procedures, tactics; and learning ability). In constructing the index, outcomes were scored as follows:

Completed (codes C1,2)=1; Resigned while performing unsatisfactorily or terminated due to inadequate report writing skills or other knowledge, skills and abilities (codes R2,7 & T2)=0.

A total of 423 students were identified as having one of these outcomes, of whom 98.3% were successful in completing basic training. The average time between Read/Write testing and academy data collection was approximately 22 months.¹⁷

Validity Evidence

Reading & Writing Test total scores and Essay Test scores were found to be significantly predictive of overall academy success/failure as measured by the above described index (r=.14 and .21, respectively). As seen in Table 16, all of the individual Reading & Writing subtests, except one (Clarity) were found to be predictive of overall academy success/failure; significant validities ranged from .09, for multiple-choice Reading Comprehension, to .13 for both the Spelling and Cloze Tests. The obtained significant validities are noteworthy in view of the high base rate for successful performance (98.3%).

¹⁷Average difference between academy ending date and Reading & Writing Test date. Subjects were excluded from the analysis if they were not tested at least 90 days prior to the end of basic training.

Table 14Descriptive StatisticsReading & Writing Test and Essay Test Scores(Academy Success/Failure Sample)

	Total R/W	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay
Mean	51.83	54.36	48.81	74.78	77.00	74.47	76.74	67.80	8.27
SD	10.34	10.32	10.46	13.70	14.72	14.01	14.67	12.41	1.59

Note: Read/Write Test N=423; Essay N=295. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores; Essay score is a raw sum of two ratings.

Basic Academy Outcome	Freq	Pct
Graduated:		
(C1) Normal time	397	85.0%
(C2) Extra time	19	4.1%
Subtotal:	416	89.1%
Failed to Complete Training:		
(R1) ResignedSatisfactory performance	14	3.0%
(R2) ResignedInadequate report writing ^a	3	0.6%
(R4) ResignedInadequate weaponless defense ^b	1	0.2%
(R7) ResignedInadequate knowledge, skill, ability other than writing, analytical or physical ^e	3	0.6%
(R8) ResignedInadequate for other reasons	17	3.6%
(T2) TerminatedInadequate report writing	1	0.2%
(T4) TerminatedInadequate weaponless defense	2	0.4%
(T8) TerminatedInadequate for other reasons	1	0.2%
(11) Injury	2	0.4%
Subtotal:	44	9.4%
Recycled/Injured:		
(N1) Injury or illness	5	1.1%
(N6) Failed physical conditioning or WSTB	1	0.2%
(N8) Other reasons	1	0.2%
Subtotal:	7	1.5%
TOTAL:	467	100%

Table 15Frequency Distribution of Basic Academy Outcomes

^aOne student was also identified as performing unsatisfactorily due to inadequate knowledge, skill or ability other than writing, analytical or physical.

^bThis student was also identified as having inadequate baton skills.

^cOne student was also identified as performing unsatisfactorily due to inadequate weaponless defense and baton skills.

Table 16Zero-Order CorrelationsReading & Writing Test and Essay Test Scores with
Academy Success/Failure

Test:	Pass/ Fail Indexª
R/W Total	.14**
Read STD	.12**
Write STD	.12**
Clarity	(.04)
Spelling	.13**
Vocabulary	.10*
M/C Reading	.09*
Cloze	.13**
Essay	.21**

**p<.01; *p<.05 (one-tailed).

Note: Read/Write Test N=423; Essay Test N=295. Correlations are point-biserials.

^aDichotomous criterion scored as follows: Graduated (C1,2)=1; Resigned or terminated due to inadequate report writing skills or other KSAs (R2,7 & T2)=0; others excluded from analysis.

Prediction of Academy Proficiency Test Scores

Descriptive Statistics

Predictors. Reading & Writing Test and Essay Test means and SDs for the Proficiency Test analysis sample are shown in Table 17.

Criterion. Table 18 contains Proficiency Test score means and SDs obtained by examinees in the validity analyses. The majority of Proficiency Test scores were obtained on form 7 (approximately 80%), followed by form 4 (approximately 15%); and less than 5% of the examinees took forms 5, 6 or 8. Reliability estimates (KR-20) for forms 4-8 were reported in a previous POST study: r_{yy} =.89, .77, .81, .77 and .76, respectively (Weiner & Berner, 1987).

The Proficiency Test was administered approximately 12 and 1/2 months after the Reading & Writing Test, on average.

Validity Evidence

As seen in Table 19, Reading & Writing total and Essay Test scores were found to be significantly predictive of subsequent Academy Proficiency Test scores (r=.47 and .28, respectively). Subtest validities ranged from .14 (Spelling) to .45 (multiple-choice Reading Comprehension). As would be expected, given that the Proficiency Test is a written examination that requires examinees to read and select the appropriate answers to multiple-choice questions, reading ability tests were more highly correlated with Proficiency Test scores than were tests of writing ability.

Table 17Descriptive StatisticsReading & Writing Test and Essay Test Scores

(Academy Proficiency Test Sample)

	Total R/W	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay
Mean	52.56	55.14	49.31	74.51	77.20	76.05	77.16	69.24	8.27
SD	9.67	9.51	10.08	13.37	14.96	13.63	13.77	11.91	1.60

Note: Read/Write Test N=13,347; Essay N=227. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores; Essay score is a raw sum of two ratings.

Table 18Descriptive StatisticsAcademy Proficiency Test Scores

	Sample 1 (Read/Write test)	Sample 2 (Essay test)
Mean	50.96	51.20
SD	10.32	10.08
N	13,347	227

Table 19Zero-Order CorrelationsReading & Writing Test and Essay Test Scores with
Academy Proficiency Test Scores

an ta san an eg

- 27

Test:	Proficiency Test
R/W Total	.47***
Read STD	.49***
Write STD	.34***
Clarity	.25***
Spelling	.14***
Vocabulary	.36***
M/C Reading	.45***
Cloze	.40***
Essay	.28***

***p<.0001 (one-tailed).

ł

Note: Read/Write Test N=13,347; Essay test N=227.

Prediction of Field Training Officer (FTO) Ratings

Descriptive Statistics

Predictors. Table 20 contains means and SDs on the Reading & Writing Tests and the Essay Test for the FTO rating sample.

Criterion. Means and SDs are shown in Table 21 for FTO ratings on each of the four writing ability scales, the four job knowledge scales, and the learning ability scale, as well as for a writing scale composite and knowledge/learning scale composite. Intercorrelations between the scales are shown in Table 22.¹⁸

Characteristics of the FTOs who evaluated the officers in the study are shown in Table 23 (note that evaluators could not be identified for some of the ratees in the study). It is noteworthy that approximately one-third of the evaluators were racial/ethnic minorities (Asians, Blacks, Hispanics, and other non-Whites) and approximately 10% were females. Of the 221 FTOs described in the table, 19% rated two or more officers.

On average, the time between the administration of the Reading & Writing Tests and completion of the FTO ratings was approximately 17 and 3/4 months.¹⁹

The 5-factor solution yielded by the Principal Components analysis accounted for 73% of the total variance. Factor names and scales loading highest on each factor are as follows: (1) Knowledge: Legal knowledge, Knowledge of dept. policy & procedure, Knowledge of patrol procedure, Knowledge of investigation; (2) General Cognitive Ability: Judgment, Observation Skills, Learning, Oral Communication, Initiative; (3) Writing Ability: Organization & Narrative, Mechanics, Information & Elements, Timeliness; (4) Tratts: Interpersonal Behavior, Teamwork, Emotional Control, Adaptability; and (5) Physical Fitness/Appearance: Appearance, Physical Fitness. The following scales loaded on multiple factors: Assertiveness (factors 1 & 2), Dependability (factors 2, 4 & 5), and Officer Safety (factors 1, 2 & 5).

The results support the use of writing composite and job knowledge composite scores. While the Learning scale was found to load highest with other cognitive abilities (which is a reasonable result), it was nevertheless included in the Job Knowledge composite in keeping with a priori hypotheses. The significant positive correlations of the Learning scale with the knowledge scale ratings support this approach.

¹⁹N=291. Average difference between FTO rating date and Reading & Writing Test date.

¹⁸Exploratory factor analyses were conducted on the four writing ability scales and 18 job element scales covering other patrol knowledge areas, abilities and personality traits and characteristics (N=300). Both Principal Components and Principal Factors algorithms were employed with Varimax rotation, and solutions were yielded suggesting 4 or 5 factors, although the factor patterns yielded by the Principal Components analysis were somewhat more interpretable. It is noteworthy that in both sets of analyses, the 4 writing scales coalesced into a distinct factor.

	Table 20	
	Descriptive Statistics	
Reading &	b Writing Test and Essay Test Scores	
	(FTO Rating Sample)	

	Total R/W	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay
Mean	51.93	54.20	49.15	74.94	77.25	75.08	76.62	67.61	8.22
SD	10.02	10.14	10.17	13.78	14.32	13.73	15.59	11.71	1.61

Note: Read/Write Test N=329; Essay N=292. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores; Essay score is a raw sum of two ratings.

	distant and the second		
	Mean	SD	Ν
Mean Writing Ability	3.78	0.64	329
1. Organization & Narrative	3.89	0.72	329
2. Writing Mechanics	3.87	0.73	329
3. Information & Elements	3.84	0.71	329
4. Timeliness	3.52	0.79	329
Job Knowledge Rating Composite ^a	3.49	0.54	329
3. Learning	3.73	0.69	328
15. Legal Knowledge	3.45	0.62	329
16. Knowledge of Dept. Policy	3.41	0.60	328
17. Knowledge of Patrol Procedures	3.46	0.61	327
18. Knowledge of Investigative Proc.	3.41	0.62	326

Table 21Descriptive Statistics for FTO Ratings

Note: Job performance rating scale: 5=Excellent, 4=Good, 3=Adequate, 2=Poor, 1=Very Poor.

^aMean of knowledge scales 15-18 and Learning scale.

	Α	В	C of	D ·	Е	F	G	Н	Ι	J	к	L
(A) Mean Writing Ability								-				
(B) Org. & Narrative	89		an An An				:	n in				
(C) Writing Mechanics	90	79			5.	a se a	a ¹					
(D) Info. & Elements	87	71	73									
(E) Timeliness	82	61	61	60								
(F) Learning	50	47	45	39	43							
(G) Legal Knowledge	59	50	49	48	58	52						
(H) Know. Dept. Policy	53	45	39	44	54	52	79					
(I) Know. Patrol Proc.	50	42	38	42	52	53	71	79				
(J) Know. Invest. Proc.	55	46	46	47	51	53	69	75	76			
(K) Know. Composite ^a	63	54	51	52	61	74	87	.90	89 [°]	87		
(L) Global Rating	59	48	49	52	55	59	57	50	58	54	65	<u> </u>

Table 22FTO Rating Intercorrelations

Note: N=317 to 329. Correlations are reported with decimals omitted; all coefficients are significant (p<.0001).

"Mean of knowledge scales G-J and Learning scale.

Agency	Freq	Pct
Oakland	6	2.7%
Los Angeles	82	37.1%
Sacramento	6	2.7%
San Diego	82	37.1%
San Francisco	45	20.4%
Race/Ethnicity		
Asian	13	5.9%
Black	19	8.6%
Hispanic	32	14.5%
White	144	65.2%
Other	13	5.9%
Gender ^a		
Male	196	90.3%
Female	21	9.7%
Total ^b	221	

Table 23Characteristics of the FTO (Rater) Sample

and a second state of the second state of the

Years of Experience	N	Mean	SD
FTO	141	3.8	.4.5
Sergeant	5	8.5	5.5
Total law enforcement	169	9.7	5.7

"Not reported for 4 raters.

5.02

^bTotal number of raters identified for 280 officers.

Validity Evidence

Table 24 contains correlation coefficients obtained for Reading & Writing Test and Essay Test scores with FTO ratings of officers' performance in field training. All tests except the Essay were found to be significantly predictive of overall writing ability (mean of four ability ratings). Essay scores were found to predict only ratings of Organization & Narrative and Writing Mechanics (r=.12 in both instances). The Spelling and Cloze tests were most highly correlated with overall writing ability (.26 and .23, respectively). Of the four writing abilities rated, the tests were generally found to best predict Writing Mechanics, while correlations with Timeliness were the lowest in magnitude.

Read/Write total scores were found to predict Job Knowledge Composite ratings and four of the five knowledge/ability component scales (knowledge of patrol procedures was not predicted). Of the individual tests, only the Spelling and Cloze Tests were significantly predictive of Knowledge Composite ratings. These two tests were the most consistently significant individual predictors of the five knowledge/ability component scales.

		Reading & Writing Test											
FTO Rating	R/W Total	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay				
Mean Writing Ability	.26***	.21***	.24***	.13**	.26***	.14**	.15**	.23***	(.09)				
1. Org & Narrative	.24***	.20***	.22***	.13**	.23***	.12*	.14**	.21***	.12*				
2. Writing Mechanics	.31***	.24***	.31***	.19***	.30***	.19***	.17**	.26***	.12*				
3. Information & Elements	.21***	.18***	.18***	.11*	.19***	.09*	.13**	.19***	(.03)				
4. Timeliness	.14**	.11*`	.14**	(.03)	.20***	(.08)	(.08)	.13*	(.05)				
Knowledge Composite [®]	.12*	.11*	.09*	(.04)	.14**	(.02)	(.08)	.12*	(.04)				
3. Learning	.13*	.14**	(.08)	(.01)	.11*	(.06)	.14**	(.09)	(.08)				
15. Legal Know.	.10*	.10*	(.08)	(.00)	.12*	(.03)	(.06)	.11*	(01)				
16. Know. Dept Policy	.10*	(.08)	.10*	(.06)	.13**	(.03)	(.05)	.11*	(.00)				
17. Know. Patrol Procedures	(.06)	(.05)	(.06)	(.06)	.10*	(02)	(.02)	(.07)	(.08)				
 Know. Investigative Procedures 	.10*	.10*	(.07)	(.03)	.11*	(01)	(.05)	.13**	(.05)				

Table 24Zero-Order CorrelationsReading & Writing Test and Essay Test Scores with FTO Ratings

***p<.001; **p<.01, *p<.05 (one-tailed).

Note: Read/Write Test N=317 to 329; Essay test N=281 to 292.

*Mean of knowledge scales 15-18 and Learning scale.

40

Prediction of Field Training Success/Failure

Descriptive Statistics

Predictors. Means and SDs on the Reading & Writing Test and Essay Test for the field training success/failure analysis sample are shown in Table 25 (i.e., examinees for whom Index #2, below, was computed).

Criterion. Table 26 contains a frequency distribution of field training outcomes that were obtained for 422 officers.²⁰ Here it is seen that 90.5% of the officers were successful in completing field training and, of the 10% who failed to complete field training, only two officers were identified as having failed for reasons related to inadequate report writing skills; and only one officer resigned while performing unsatisfactorily due to inadequate analytical skills. Thus, there was little variance in this criterion measure of field training performance that was attributed directly to inadequate writing skills.

Two dichotomous pass/fail indices were constructed to quantify the various field training outcomes for purposes of conducting validity analyses in a manner similar to that described in the above analysis of basic academy success/failure. Again, the aim was to identify outcomes that were conceivably relevant to reading and writing abilities. The indices included:

Index #1: Completed (C1-3)=1, Resigned or terminated due to inadequate report writing, analytical skills, or other knowledge, skills and abilities (R3,4,5 and T3)=0.

Index #2: Completed (C1-3)=1; Resigned or terminated due to inadequate report writing, analytical skills, or other knowledge, skills and abilities, or performance level unknown (R3,4,5,7 and T3,7)=0.

A total of 388 officers were identified as having completed or failed to complete field training for reasons included in Index #1; 403 officers completed or failed to complete field training for reasons included in Index #2. The success rates²¹ for the two indices were 98.5% and 94.8%, respectively. The average time between Read/Write testing and field

²¹Number of officers completing field training divided by the sum of the number completing plus the number failing to complete field training.

 $^{^{20}}$ For seven of the officers, two outcomes were coded. In six of these cases, the officers were initially injured, ill, or on leave (codes O1 and 2) but did eventually complete field training (codes C1, 2 and 3) -- these cases were included with the "Completed" group. In the remaining case, the officer apparently resigned while performing unsatisfactorily due to inadequate knowledge and skills other than physical, writing or analytical (code R5), but the officer continued in a non-patrol assignment (code F5) -- this case was included with the "Failed to Complete" group and was counted once in any pass/fail index that included the R5 or F5 code.

Table 25Descriptive StatisticsReading & Writing Test and Essay Test Scores(Field Training Success/Failure Sample)

	Total R/W	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay
Mean	51.27	53.47	48.70	74.00	77.53	74.39	75.15	67.31	8.17
SD	10.43	10.69	10.24	13.59	14.62	13.71	16.31	12.12	1.62

Note: Read/Write Test N=403; Essay N=336. Total Read/Write, Reading STD and Writing STD scores are T-scores calibrated to 1983 validation study sample; subtest scores are raw percent correct scores; Essay score is a raw sum of two ratings.

12.

4 * *

Field Training Outcome	Freq	Pct
Completed:		
(C1) Normal time	363	86.0%
(C2) Extra time (remediation) ^a	17	4.0%
(C3) Time required unknown ^b	2	0.5%
Subtotal:	382	90.5%
Failed to Complete Field Training:		
(R1) ResignedSatisfactory performance	6	1.4%
(R3) ResignedInadequate report writing skills	1	0.2%
(R4) ResignedInadequate analytical skills	1	0.2%
(R5) ResignedInadequate KSA other than physical, writing, or analytical ^c	3	0.9%
(R6) ResignedOther reasons	2	0.5%
(R7) ResignedPerformance level unknown	14	3.3%
(T3) TerminatedInadequate report writing skills	- 1	0.2%
(T7) TerminatedPerformance level unknown	- 1	0.2%
(O1) Injųry or illness	4	0.9%
(O2) Other (military reserve, transferred, etc.)	7	1.7%
Subtotal:	40	9.5%
TOTAL:	422	

Table 26Frequency Distribution of Field Training Outcomes

*Five of these officers were initially injured, ill or on leave, but eventually completed training.

^bOne officer was initially injured, ill or on leave, but eventually completed training.

One officer was reportedly assigned to a non-patrol position.

training data collection was approximately 22 months for those who completed field training and 19 months for the small number of subjects who failed to complete training.²²

Validity Evidence

No significant relationships were detected between Reading & Writing Test and Essay Test scores and the two indices of field training success/failure (p>.05, one-tailed). These results are shown in Table 27. The lack of significant findings is not surprising in view of the very low rate of failure reported for reasons related to reading and writing ability. Again, there was very little variance in the criterion index of success/failure to be predicted by test scores.

²²Average time between testing and either field training completion date or separation date. Test date and field training completion date were available for 265 subjects; test date and agency separation date were available for 17 subjects.

Table 27Zero-Order CorrelationsReading & Writing Test and Essay Test Scores with
Field Training Success/Failure

Test:	Field Training Success/Failure					
	Index #1	Index #2				
R/W Total	01	.05				
Read STD	.02	.06				
Write STD	03	.04				
Clarity	.01	.08				
Spelling	04	01				
Vocabulary	04	.02				
M/C Reading	.03	.07				
Cloze	.00	.03				
Essay	07	.05				

Note: Correlations are point-biserials and are *non*-significant (p>.05, one-tailed). Read/Write Test N=388 for Index #1, N=403 for Index #2; Essay Test N=323 for Index #1, N=336 for Index #2.

Success/failure indices:

Index #1: Completed (C1-3)=1; Resigned or terminated due to inadequate report writing, analytical skills, or other knowledge, skills and abilities (R3,4,5 and T3) = 0.

Index #2: Completed (C1-3)=1; Resigned or terminated due to inadequate report writing, analytical skills, or other KSAs, or performance level unknown (R3,4,5,7 and T3,7) = 0.

Differential Prediction Analysis

A statistical procedure was conducted to examine the extent to which Reading & Writing Test total scores are equally predictive of academy and field performance for racial/ethnic and gender minority and majority groups. The analysis was also conducted for Essay Test scores. A separate analysis was conducted comparing each racial/ethnic and gender minority group to the corresponding majority group²³ in the following three steps:

(1) minority and majority group error variances resulting from the overall regression of criterion scores onto test scores were compared (F SE) and if a statistically significant difference was detected (p<.05), the analysis was stopped at this point as differential prediction would be evidenced; (2) if error variance differences were not detected, then the slopes of the separate criterion-test score regression lines for majority and minority groups were compared (t slopes); and (3) if no statistically significant slope differences were detected, then the intercepts of the separate criterion-test score regression lines for majority and minority groups were compared (t slopes); and (3) if no statistically significant slope differences were detected, then the intercepts of the separate criterion-test score regression lines for the separate criterion-test score regression lines were compared (t intercepts).

In addition to the above comparison of regression parameters, the mean difference was computed between each minority group's actual criterion performance and that predicted on the basis of the overall regression of criterion scores onto test scores (**Residuals mean**), and a statistical significance test (t-Res) was performed to determine whether the residual was significantly greater or less than zero. A negative residual indicates that test scores *over*predict criterion performance; that is, test scores depict the minority group's performance more favorably than the actual criterion measure of academy/field performance. A positive residual means just the opposite.

This analysis was modeled after the methodological approach to investigating predictive bias described in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1985, pp. 12-13) and is consistent with the definition of fairness espoused by federal guidelines pertaining to employment testing [cf. *Uniform Guidelines on Employee Selection Procedures*, EEOC, et al., 1978, sec. 14B(8a)].

Academy Instructor Ratings

Results of the analysis conducted for Read/Write total scores predicting academy instructor ratings of students' writing ability demonstrated in basic training are presented in Table 28. Results for Essay Test score predictions of the same criterion are shown in Table 29. The tables also contain within-race and within-sex descriptive statistics and validities.

The results indicate that Reading & Writing total scores were significantly predictive of academy ratings for all subgroups studied (i.e., Asians, Blacks, Hispanics, Whites, males

46

²³Whites comprised the racial/ethnic majority group, with whom American Indians, Asians, Blacks, Filipinos, and Hispanics (minority groups) were compared. Males comprised the gender majority group.

Table 28

Differential Prediction Analysis by Race/Ethnicity and Gender Total Read/Write Score Predicting Academy Instructor Rating

			Descriptive	e Statistics		. <u> </u>
	Asian	Black	Hispanic	White	Female	Male
Test Battery Score N Mean SD	29 50.70 10.60	71 45.33 8.93	86 46.72 10.17	314 55.03 9.86	147 51.80 10.28	373 51.81 10.85
Criterion Score N Mean SD r (with test)	28 3.64 0.56 .44*	68 3.38 0.71 .30**	86 3.50 0.66 .31**	303 3.82 0.55 .32***	145 3.84 0.60 .36***	356 3.62 0.63 .43***
		T	ests of Regress	sion Paramete	ers	· · · ·
F SE t slopes t intercepts	1.04 -0.53 0.97	1.69** N/A N/A	1.49** N/A N/A	N/A	1.02 0.73 -4.20***	N/A
Residuals Mean t-Res	-0.09 -0.94	-0.17 -2.10*	-0.11 -1.66	N/A	0.17 3.57***	N/A

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: residual = actual minus predicted criterion scores.

Table 29Differential Prediction Analysis by Race/Ethnicity and Gender
Essay Test Score Predicting Academy Instructor Rating

	Descriptive Statistics							
	Asian	Black	Hispanic	White	Female	Male		
Essay Score N Mean SD	23 8.13 1.74	53 7.74 1.67	74 7.85 1.76	250 8.43 1.53	123 8.44 1.69	289 8.09 1.59		
Criterion Score N Mean SD r (with test)	23 3.60 0.58 (.18)	53 3.38 0.72 .23*	74 3.53 0.67 .44***	250 3.79 0.51 .23***	123 3.85 0.56 .31***	289 3.60 0.60 .29***		
	Tests of Regression Parameters							
F SE t slopes t intercepts	1.35 0.29 1.47	2.00** N/A N/A	1.51* N/A N/A	N/A	1.15 0.23 -3.55**	N/A		
Residuals Mean t-Res	-0.15 -1.23	-0.28 -2.92**	-0.15 -2.08*	N/A	0.15 3.11**	N/A		

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: residual = actual minus predicted criterion scores.

and females). Essay scores were found to predict academy ratings for all groups except one (a positive, but nonsignificant correlation was obtained for Asians).²⁴

Differential prediction of academy ratings was detected for both the Read/Write and Essay Tests in a similar manner; i.e., for Blacks (SE), Hispanics (SE), and females (intercepts), but not Asians. The net result of these differences was neutral for Asians and Hispanics (Read/Write scores) in that their performance was neither over- or underpredicted; i.e. their mean residuals were not significantly different from zero. The net impact for Blacks and Hispanics (Essay scores) was actually favorable as their performance was *over* predicted. The results by gender were somewhat different, however, in that females' academy performance was significantly *under* predicted by both Read/Write total scores and Essay scores.

Academy Success/Failure

Regression parameters for racial/ethnic and gender majority and minority groups were also compared with respect to test score predictions of overall basic academy success/failure. The results in Table 30 indicate that Reading & Writing Test total scores were significantly predictive of academy success/failure only for the larger samples; i.e., Whites and males. Validity coefficients obtained for racial/ethnic minorities and females were *not* statistically significant, although positive correlations were obtained for Blacks, Hispanics and females (correlations obtained for the latter two groups were of the same magnitude as those obtained for Whites and males). Due to the relatively small sample sizes obtained for the minority groups, statistical power was low to detect a significant relationship of the magnitude obtained for the total sample.²⁵

Differential prediction of academy success/failure was evidenced for Asians, Blacks, Hispanics and females; i.e., significant differences in regression error variances were identified for racial/ethnic minorities relative to Whites and for females relative to males. However, the net effect of this result was negligible since test score predictions of each group's academy success/failure were consistent with actual success/failure, on average. That is, the mean residuals were approximately zero.

The results in Table 31 indicate that Essay Test scores were predictive of academy success/failure for Hispanics, Whites, females and males. The correlation obtained for Blacks was positive and of the same magnitude as that obtained for Whites, but was nonsignificant due to the small sample size (N=39) (Asians were not studied because all of the subjects completed basic training; i.e., there was no criterion variance for this group). Differential prediction of academy success was found for Blacks (SE) and females (SE), but not Hispanics. The mean residuals for these three groups were zero, indicating that their academy success/failure was *not under*predicted by Essay scores.

²⁴The 95% confidence interval for the correlation of .18 obtained for Asians ranges from approximately -.23 to .59.

²³For example, in order to have 80% power to detect significance at the .05 level when r=.20, a sample size of 160 is needed (Cohen, 1988).

Differential	Predictio	n Analysi	s by Rac	e/Ethnicit	y and Gende	er
Total Reac	l/Write Sc	ore Predic	cting Ac	ademy Su	ccess/Failure	•
		1				

Table 30

	Descriptive Statistics							
	Asian	Black	Hispanic	White	Female	Male		
Test Battery Score N Mean SD	27 48.16 11.07	58 45.97 8.88	64 48.46 9.02	258 54.48 9.95	79 52.76 9.80	342 51.63 10.45		
Criterion Score N Mean SD r (with test)	27 .963 .193 (.00)	58 .966 .184 (.07)	64 .969 .175 (.16)	258 .992 .088 .16**	79 .987 .113 (.15)	342 .983 .132 .13**		
	Tests of Regression Parameters							
F SE t slopes t intercepts	4.95*** N/A N/A	4.48*** N/A N/A	3.99*** N/A N/A	N/A	1.37* N/A N/A	N/A		
Residuals Mean t-Res	-0.02 -0.50	-0.01 -0.45	-0.01 -0.46	N/A	0.00 0.19	N/A		

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: Academy success/failure index scored as follows: Graduated (C1,2)=1; Resigned or terminated due to inadequate report writing skills or other KSAs (R2,7 & T2)=0; others excluded from analysis. Residual = actual minus predicted criterion scores.

	Descriptive Statistics						
	Black Hispanic		White	Female	Male		
Essay Score							
N	29	46	144	62	233		
Mean	7.69	8.30	8.53	8.74	8.15		
SD	1.51	1.72	1.51	1.61	1.57		
Criterion Score		· · ·					
N	29	46	144	62	233		
Mean	.966	.978	.986	.984	.979		
SD	.186	.147	.117	.127	.145		
r (with test)	(.21)	.38**	.20**	.38**	.17**		
	Tests of Regression Parameters						
F SE	2.50**	1.43		1.45*			
t slopes	N/A	-1.34	N/A	N/A	N/A		
t intercepts	N/A	0.15		N/A			
Residuals							
Mean	0.00	0.00	N/A	.0.00	N/A		
t-Res	-0.14	-0.12		-0.29			

Table 31Differential Prediction Analysis by Race/Ethnicity and GenderEssay Test Score Predicting Academy Success/Failure

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: Academy success/failure index scored as follows: Graduated (C1,2)=1; Resigned or terminated due to inadequate report writing skills or other KSAs (R2,7 & T2)=0; others excluded from analysis. Residual = actual minus predicted criterion scores.

Academy Proficiency Test Scores

Read/Write Test score predictions of Academy Proficiency Test performance were compared for several racial/ethnic minorities (American Indians, Asians, Blacks, Filipinos, and Hispanics) versus Whites, as well as between gender groups. The results in Table 32 indicate that Read/Write total scores were significantly predictive of Proficiency Test scores for all subgroups studied, with validities ranging from .40 (Blacks and Hispanics) to .57 (American Indians). While significant differences in regression parameters were detected for American Indians (slopes), Asians (SE), Blacks (SE), Filipinos (intercepts), Hispanics (SE), and females (SE), the net effect of such differences was neutral for American Indians (i.e., the mean residual was not significantly greater than zero) and favorable for the remaining groups. That is, Proficiency Test scores were significantly *over*predicted by Read/Write total scores for Asians, Blacks, Filipinos, Hispanics and females.

Essay Test score predictions of Proficiency Test scores were examined for only two racial/ethnic minorities (Blacks and Hispanics) due to the small number of examinees for whom Essay scores were available. Gender group comparisons were also made. As seen in Table 33, Essay scores were significantly predictive within sex but *not* within race (only Whites' Proficiency Test scores were predicted by Essay scores). The nonsignificant findings for Blacks and Hispanics are inconclusive due to the small sample sizes for these groups. Differential prediction was detected for Blacks (slopes), Hispanics (intercepts), and females (SE). However, all minority groups' Proficiency Test scores were significantly *over*predicted by Essay scores, on average.

FTO Ratings

Table 34 contains results of racial/ethnic and gender subgroup analyses with respect to Read/Write total score predictions of FTO ratings of officers' writing abilities demonstrated throughout field training. As seen in the table, significant test-criterion correlations were obtained for Whites, males, and females, while the correlations obtained for Asians, Blacks and Hispanics were positive but nonsignificant.²⁶

Evidence of differential prediction by Read/Write total scores was found for Asians (SE), Blacks (SE) and females (intercepts), but not Hispanics. None of the racial/ethnic minority groups' performance ratings were underpredicted (residuals were near zero for all such groups). However, it is noteworthy that females' rated writing abilities were significantly *under*predicted, on average.

 $^{^{26}}$ Again, the small sample sizes for these groups may be a contributing factor. For example, consider the confidence interval for a sample of 50 and a correlation of .10. In this case there is a 95% probability that the true value of the correlation for the total population lies between -.19 and +.39. Thus, when statistical significance is not detected for a small sample, there is little in the way of conclusions that may be drawn.

Table 32

Differential Prediction Analysis by Race/Ethnicity and Gender Total Read/Write Score Predicting Academy Proficiency Test Score

-	r									
		Descriptive Statistics								
	Am. Indian	Asian	Black	Filipino	Hispanic	White	Female	Male		
Test Battery Score N Mean SD	137 50.96 9.49	407 51.16 10.19	996 48.43 9.99	154 50.33 10 10	1821 49.00 9.69	9495 53.85 9.26	1882 54.20 9.11	11,149 52.35 9.69		
Criterion Score N Mean SD r (with test)	137 51.90 11.42 .57***	407 49.37 9.63 .52***	996 45.26 10.30 .40***	154 47.23 10.21 .43***	1821 47.46 10.31 .40***	9495 52.37 9.95 .45***	1882 48.06 10.83 .49***	11,149 51.47 10.15 .48***		
		Tests of Regression Parameters								
F SE t slopes t intercepts	1.17 -2.42* N/A	1.17** N/A N/A	1.14*** N/A N/A	1.07 0.62 4.75***	1.14*** N/A N/A	N/A	1.13*** N/A N/A	N/A		
Residuals Mean t-Res	0.92 1.12	-1.62 -3.97***	-3.98 -13.26***	-3.38 -4.55***	-2.12 -9.53***	N/A	-3.71 -16.99***	N/A		

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: Residual = actual minus predicted criterion scores.

Table 33

Differential Prediction Analysis by Race/Ethnicity and Gender Essay Test Score Predicting Academy Proficiency Test Score

		Descriptive Statistics						
	Black Hispanic White		Female	Male				
Essay Score N Mean SD	26 7.38 1.83	36 8.22 1.76	142 8.47 1.44	48 8.48 1.66	179 8.22 1.58			
Criterion Score N Mean SD r (with test)	26 40.71 8.80 (09)	36 45.84 8.24 (.06)	142 54.98 8.30 .29***	48 47.68 12.09 .25*	179 52.14 9.19 .32***			
		Tests of Regression Parameters						
F SE t slopes t intercepts	1.52 2.09* N/A	1.14 1.53 5.83***	N/A	1.80** N/A N/A	N/A			
Residuals Mean t-Res	-10.24 -5.32***	-7.00 -4.95***	N/A	-3.88 -2.30*	N/A			

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: Residual = actual minus predicted criterion scores.

	Descriptive Statistics							
	Asian	Black	Hispanic	White	Female	Male		
Test Battery Score N Mean SD	30 51.45 11.20	47 46.04 9.46	49 47.70 9.75	205 54.67 9.04	93 51.76 9.86	236 51.99 10.11		
Criterion Score N Mean SD r (with test)	21 3.79 0.75 (.12)	47 3.51 0.73 (.07)	49 3.66 0.59 (.17)	205 3.88 0.59 .25***	93 4.01 0.60 .20*	236 3.69 0.63 .29***		
	Tests of Regression Parameters							
F SE t slopes t intercepts	1.72* N/A N/A	1.68* N/A N/A	1.05 0.64 1.28	N/A	1.07 0.75 -4.30***	N/A		
Residuals Mean t-Res	-0.02 0.06	-0.18 -1.71	-0.09 -1.08	N/A	0.23 3.74***	N/A		

Table 34Differential Prediction Analysis by Race/Ethnicity and Gender
Total Read/Write Score Predicting FTO Ratings

***p<.001; **p<.01; *p<.05 (one-tailed for r; two-tailed for t residuals).

Note: Criterion is mean rating across four writing ability scales. Residual = actual minus predicted criterion score.
Results for the Essay Test are shown in Table 35. The criterion in this case was the mean of two writing abilities, Organization & Narrative and Writing Mechanics, since significant correlations were obtained only for these ratings. With regard to within-group validities, only males' performance ratings were significantly predicted, although positive correlations of comparable magnitude were obtained for all racial/ethnic groups. Differential prediction of rated writing ability was detected for Blacks (SE), Hispanics (intercepts) and females (intercepts). Blacks' and Hispanics' performance ratings were significantly *under*predicted by Essay scores, while females' performance was significantly *under*predicted by Essay scores.

Summary of Validity Evidence

Overall Results

Reading & Writing Test total scores and subtest scores were found to be significantly predictive of *academy performance* as measured by instructor ratings, Proficiency Test scores, and with one exception (Clarity), overall success/failure in completing basic training. Total and subtest scores were also predictive of performance in *field training* as measured by FTO ratings of writing ability, but were not found to predict overall field training success/failure.

Essay Test scores were found to predict *academy performance* as measured by instructor ratings of demonstrated writing ability, overall academy success/failure, and Academy Proficiency Test scores. Essay scores were found to predict only certain aspects of performance in *field training*; i.e., FTO ratings of two writing elements (Organization & Narrative and Mechanics) were predicted. No significant correlations were obtained between Essay scores and *overall field training success/failure* indices.

Table 36 summarizes the validity results obtained in the present study. Adjusted validity coefficients are included in the table in those instances where a significant zero-order correlation was obtained. These adjusted correlations take into account statistical artifacts that are known to reduce the value of obtained correlations, including: (1) range-restriction in test scores, and (2) for academy ratings, FTO ratings, and Proficiency Test scores only, attenuation in the criterion due to unreliability. More specifically, corrections were made to account for the reduced variation in test scores obtained by subjects in the study since they were hired, at least in part, on the basis of their scores on the POST Read/Write Test or other tests that are likely correlated with the POST test (the earlier reported applicant SDs were used in this correction). A second adjustment was made to correlations with performance ratings and Proficiency Test scores to account for unreliability in these measures. The earlier reported interrater reliability estimate (.77) that was derived for academy instructor ratings; a weighted mean of Proficiency Test form reliabilities (.789) was used to adjust correlations

		Descriptive Statistics							
	Black	Hispanic	White	Female	Male				
Essay Score			-		· ·				
N	39	46	182	83	209				
Mean	7.90	8.00	8.41	8.64	8.06				
SD	1.71	1.66	1.53	1.64	1.56				
Criterion Score									
N	39	46	182	83	209				
Mean	3.58	3.70	3.99	4.12	3.78				
SD	0.76	0.62	0.62	0.64	0.66				
r (with test)	(.12)	(.17)	(.10)	(04)	.14*				
		Tests of R	egression Pa	rameters					
F SE	1.48*	1.02		1.03					
t slopes	N/A	-0.38	N/A	1.46	N/A				
t intercepts	N/A	2.68**		-3.73***					
Residuals									
Mean	-0.32	-0.22	N/A	0.22	N/A				
t-Res	-2.63*	-2.41*		3.15**					

Table 35Differential Prediction Analysis by Race/Ethnicity and GenderEssay Test Score Predicting FTO Ratings

***p<.001; **p<.01; *p<.05 (one-tailed for r, two-tailed for t residuals).

Note: Criterion is mean of ratings on two writing ability scales (Organization & Narrative and Writing Mechanics). Residual = actual minus predicted criterion score.

		Basic Acaden	ny		Field Training			
	Instructor Rating ^a	Pass/ Fail ^b	Academy Proficiency Test ^e	FTO Rating: Writing ^d	FTO Rating: Job Know ^e	Pass/ Fail ^f		
R/W Total	.51/.39	.17/.14	.65/.47	.37/.26	.17/.12	(.05)		
Reading STD	.47/.36	.15/.12	.68/.49	.29/.21	.16/.11	(.06)		
Writing STD	.45/.35	.14/.12	.45/.34	.32/.24	.12/.09	(.04)		
Clarity	.29/.24	(.04)	.31/.25	.16/.13	(.04)	(.08)		
Spelling	.37/.28	.15/.13	.18/.14	.34/.26	.19/.14	(01)		
Vocabulary	.34/.26	.12/.10	.49/.36	.19/.14	(.02)	(.02)		
M/C Read	.42/.33	.11/.09	.62/.45	.19/.15	(.08)	(.07)		
Cloze	.41/.31	.16/.13	.55/.40	.33/.23	.18/.12	(.03)		
Essay	.45/.31	.28/.21	.42/.28	(.09)	(.04)	(.05)		

 Table 36

 Summary of Validity Evidence for Reading and Writing Tests

Note: All correlations are significant (p < 0.5, one-tailed), except those in parentheses. Both adjusted and zero-order correlations are reported, except where non-significant zero-orders were obtained. Adjusted correlations are reported first and all are corrected for range-restriction in test scores (using applicant SDs reported earlier). Adjusted correlations with Academy ratings and FTO ratings are also corrected for criterion unreliability using the interrater reliability estimate of .77 reported earlier. Adjusted correlations with Proficiency Test scores are also corrected for criterion unreliability using an internal consistency estimate of .789 (weighted mean across forms). See Guilford and Fruchter (1973).

^aAcademy instructor mean rating of four writing abilities; Read/Write Test N=504; Essay Test N=413.

^bAcademy pass/fail index: Graduated (C1,2)=1; Resigned or terminated due to inadequate report writing skills or other KSAs (R2,7 & T2)=0; Read/Write Test N=423; Essay Test N=295.

Read/Write Test N=13,347; Essay Test N=227.

^dFTO mean rating of four writing abilities; Read/Write Test N=329; Essay Test N=292.

*FTO mean rating of knowledge scales 15-18 and learning scale; Read/Write Test N=329; Essay Test N=292.

^fField Training pass/fail Index #2: Completed (C1-3)=1; Resigned or terminated due to inadequate report writing, analytical skills, or other KSAs, or performance level unknown (R3,4,5,7 & T3,7)=0; Read/Write Test N=403; Essay Test N=336.

with Proficiency Test scores.²⁷ See Guilford and Fruchter (1973) for further details regarding these corrections.

Thus, the best estimates of the validity of Read/Write total scores in predicting ratings of demonstrated writing ability are .51 for performance in the basic academy and .37 for field training performance. With regard to prediction of Academy Proficiency Test performance, the best estimate of validity is .65 for Read/Write total scores.

Within-Group Validity

Read/Write total scores were found to predict *academy* performance ratings and Proficiency Test scores within all racial/ethnic and gender groups studied. However, overall academy success/failure was *not* predicted for racial/ethnic minorities or females, and *field training* performance ratings were predicted within sex, but *not* within racial/ethnic minority groups.²⁸ Essay scores were predictive of *academy* ratings within race (except Blacks) and sex, while other academy performance measures (Overall success/failure and Proficiency Test scores) were predicted within sex and some racial/ethnic groups; and *field training* performance ratings were not predicted within race/ethnicity or for females.

Differential Prediction

Overall, the results indicate that Read/Write total scores and Essay Test scores are *not* unfair to racial/ethnic minorities in predicting measures of their performance in basic training and field training. That is, their performance was not significantly *under*predicted by test scores; in fact, their performance was often *overpredicted*. The results for females were mixed, however, in that differential prediction was detected for some performance measures, sometimes resulting in *overprediction* (Proficiency Test scores) and other times resulting in *under*prediction (academy and field training performance ratings).

²⁷Mean of the five KR-20 values reported earlier, weighted by the proportion of subjects taking each form of the Proficiency Test (see section entitled "Prediction of Academy Proficiency Test Scores" for these values).

²⁸Within-group analyses were not conducted for overall field training success/failure in view of the nonsignificant overall validity coefficient obtained.

Analysis of Alternative Test Batteries

Several alternative test batteries were constructed consisting of different configurations of the five Reading & Writing subtests and the Essay Test. The aim was to explore the possibility of improving both the prediction of academy and field training performance and the relative performance of racial/ethnic and gender minority groups. The alternative batteries were assembled in consideration of several factors, including: (a) individual subtest validities, (b) the relative difficulty of the subtests for racial/ethnic and gender groups, (c) the joint relationships among the tests as predictors of the various criteria, and (d) representation of reading and writing abilities.

Individual Test Validities

Examination of the zero-order correlations reported earlier for the Reading & Writing Tests and the Essay Test with the various performance criteria indicates that comparable prediction of certain criterion measures may be achieved with fewer tests than contained in the entire Read/Write battery (see Table 36). The uncorrected validities obtained for each test are summarized below.

- Academy instructor ratings were approximately equally predicted by the multiple-choice (M/C) Reading (.33), Essay (.31), Cloze (.31), and Spelling (.28) Tests, and the correlations obtained for Vocabulary (.26) and Clarity (.24) were not much lower. The correlation obtained for Read/Write total was .39.
 - Overall academy success/failure was best predicted by the Essay (.21), followed by the Cloze (.13), Spelling (.13), Vocabulary (.10), and M/C Reading (.09) Tests. The correlation for Read/Write total was .14.
- Academy Proficiency Test scores were best predicted by the M/C Reading (.45), Cloze (.40), and Vocabulary (.36) Tests; the Essay (.28), Clarity (.25), and Spelling (.14) Tests were also significantly predictive. The correlation for Read/Write total was .47.
 - FTO ratings of writing ability were best predicted by the Spelling (.26) and Cloze (.23) Tests; the M/C Reading (.15), Vocabulary (.14), and Clarity (.13) Tests were also significantly predictive; but Essay scores were not predictive. The correlation for Read/Write total was .26.

Relative Difficulty of Individual Tests

53

Another consideration that was made in identifying potential alternative batteries was the relative difficulty of the tests for various racial/ethnic and gender groups.

The relative difficulty of the Reading & Writing Tests and Essay Test for various groups was assessed by comparing examinees' mean standard scores between subtests within racial/ethnic and gender group. The earlier described samples of job applicants and academy students (N>120,000 Read/Write scores; N=818 Essay scores) were used for this analysis. The test scores were standardized to Z-scores within test form (200-270, 400 and 440) so that the performance of each subgroup could be compared across tests; i.e., so that the scores on each test would be in comparable score units.²⁹ These means represent the performance of examinees of all ranges of ability, including lower scoring examinees who were never hired by a law enforcement agency.

The subgroup mean Z-scores on each test are shown in Table 37. The results suggest:

- The **Spelling** Test is relatively the least difficult of the subtests for all racial/ethnic minority groups (except American Indians, although the test is not extremely difficult for this group) and for females.
- The reading tests, **M/C Reading** and **Cloze**, are relatively the most difficult of the subtests for Blacks, Hispanics and Filipinos. The **Cloze** Test is also among the most difficult for Asians.

The **Vocabulary** Test is also among the relatively most difficult tests for Hispanics and Asians.

The Essay Test is also among the relatively most difficult tests for Blacks, but is among the relatively least difficult tests for females.

 29 Z-scores are expressed in standard deviation units, so that a score of 0 is equal to the mean, a score of .10 equals one-tenth of a SD above the mean, a score of .50 equals one-half SD below the mean, and so on.

	Am. Indian	Asian	Black	Hispanic	Filipino	White	Female	Male
Total R/W	09	19	47	35	33	+.24	+.04	.00
Clarity	12	01	33	23	23	+.16	+.10	02
Spelling	- 14	+.22	11	09	+.44	+.03	+.28	06
Vocabulary	01	32	37	30	33	+.21	09	+.03
M/C Read	07	20	- 47	33	47	+.24	06	+.02
Cloze	03	33	44	34	54	+.25	03	+.02
Essay	N/A	14	-,54	15	N/A	+.20	+.24	08
N Read/Write	1,774	3,913	16,776	22,219	2,689	75,101	22,990	100,166
N Essay	्रा होता.सम् २४ २५४३ हत् रहत्वात् स्थान स्थलात् । स्था स्टब्स्	35	124	118		473	233	554

Table 37Relative Difficulty of Reading & Writing TestsMean Z-scores by Race/Ethnicity and Gender

Note: Based on earlier described sample of job applicants and academy students. Read/Write Test scores were standardized by form (200-270, 400 and 440).

Multiple Regression Analyses

A statistical procedure (multiple regression analysis) was conducted to examine the joint relationships among the individual Reading & Writing Tests and the Essay Test in predicting academy and field training performance. Major findings of the analysis are outlined below.

Academy Ratings: When the five Read/Write subtests and the Essay test were entered into a multiple regression equation predicting academy ratings, only the **Spelling**, **M/C Reading**, and **Essay** Tests were significantly predictive (i.e., received significant regression weights; p<.05, one-tailed).

Essay scores were also found to add significantly to Read/Write total scores in predicting academy ratings. However, even with optimal weighting of test scores, the multiple correlation was only .40 versus the zero-order correlation of .39 obtained for Read/Write total scores (which are computed without optimal weighting).

Overall Academy Success/Failure: Only the Essay and Spelling Tests received significant regression weights in predicting overall academy success/failure (Clarity received a *negative* weight).

When overall academy success/failure was regressed onto Essay and Read/Write total scores, only the Essay Test was found to be significantly predictive (the regression weight for Read/Write total scores was nonsignificant).

Proficiency Test: The Clarity, Vocabulary, M/C Reading, and Cloze Tests received significant regression weights in predicting *Proficiency Test* scores (Spelling received a *negative* weight and the weight for Essay scores was nonsignificant).

When Proficiency Test scores were regressed onto Essay and Read/Write total scores, only Read/Write total scores were found to be significantly predictive (the regression weight for Essay scores was nonsignificant).

FTO Ratings of Writing Ability: Only the Spelling and Cloze tests received significant regression weights among the six tests in predicting FTO ratings.

When FTO ratings were regressed onto Essay Test and Read/Write total scores, only Read/Write total scores received a significant regression weight.

The results of the above described multiple regression analyses are contained in Appendix G.

Alternative Test Batteries

The following alternative batteries were constructed in consideration of the individual test validities, the relative difficulty of the tests, the multiple regression results, and the representation of both reading and writing abilities:

- 1. **Read/Write total+Essay:** sum of Read/Write total T-score (as computed operationally) and Essay T-score; this battery represents the predictive value of adding the Essay Test to the current battery; these scores were jointly predictive of academy ratings.
- 2. Reading Composite+Essay: sum of reading composite T-score (M/C Reading + Cloze Test, as scored operationally) and Essay T-score; this battery represents the replacement of multiple-choice writing tests with the Essay; the M/C Reading and Cloze tests were jointly predictive of Proficiency Test scores.
- **3. Cloze+Essay:** sum of Cloze T-score and Essay T-score; this battery represents broad reading and writing abilities with a minimum number of tests, while eliminating a relatively difficult test for minorities (M/C Reading).
- 4. M/C Reading+Essay: sum of M/C Reading T-score and Essay T-score; same rationale as above, using an alternative reading test; these were jointly predictive of academy ratings.
- 5. Writing Composite+M/C Reading: sum of writing composite T-score (Spelling, Clarity & Vocabulary score, as computed operationally) and M/C Reading T-score; this battery was constructed in an attempt to reduce the relative difficulty for minorities while maintaining representation of reading ability.
- 6. Writing Composite+Cloze: sum of writing composite T-score (Spelling, Clarity & Vocabulary score, as computed operationally) and Cloze Test Tscore; same rationale as above, using an alternative reading test.

Validity Evidence for Alternative Batteries

Total Sample. Zero-order correlations were computed between the six alternative batteries and each of the academy and field training performance measures. The alternative batteries were found to predict academy and field training performance comparably to, or better than the current five-test Read/Write Battery. Four alternative batteries, #1-#4, yielded

statistically significant increases in validity in the prediction of *overall academy* success/failure. The magnitudes of these increases ranged from .07 to .08. See Table 38.

Overall, the results suggest that there is little or no significant gain in validity realized by adding the Essay Test to the current battery, although validity comparable to that yielded by the current battery may be achieved with fewer than five tests.

Within-Group Validity. The validity coefficients for the alternative batteries computed within race/ethnicity and gender do not suggest a single alternative to the current battery that improves prediction of performance for all groups. The within-group results are presented in Table 39 and are outlined below.

- None of the alternative batteries significantly improved prediction of *Proficiency Test* performance for any racial/ethnic or gender subgroups.
 - None of the alternative batteries were found to improve prediction of Asians' or Blacks' performance.

Prediction of *academy instructor ratings* was improved for Hispanics by Battery #1 (Read/Write total+Essay).

Prediction of *overall academy success/failure* was improved for Hispanics, Whites, males, and females by Battery #1; for Hispanics, males and females by Battery #2 (Reading Composite+Essay); for females by Battery #3 (Cloze+Essay); and for Hispanics and females by Battery #4 (M/C Reading+Essay).

- .
- Prediction of FTO Ratings was improved for Hispanics by Battery #6 (Writing Composite+Cloze).

Relative Difficulty of Alternative Batteries

The relative difficulty of the alternative test batteries for racial/ethnic and gender subgroups is summarized in Table 40. The scores were standardized to a Z-scale as described above and the sample was, again, comprised of job applicants in order to realistically portray the difficulty of the tests for examinees with a wide range of ability.³⁰ Overall, the alternative batteries were not found to offer many significant reductions in difficulty for minorities relative to the current Read/Write Battery. The only statistically significant reduction in difficulty observed for minorities was for females, with Batteries #1 and #5 (note that Battery #1 was significantly *harder* for Blacks than the current battery).

³⁰The aforementioned sample of applicants at the City of Sacramento were used for the analysis.

Table 38Validity Evidence for AlternativeReading and Writing Test Batteries

Test Battery	Academy Instructor Rating ^a	Academy Pass/Fail ^b	Academy Proficiency Test ^e	FTO Rating ⁴
1. R/W Total+Essay	.40***	<u>.22</u> ***	.50***	.20**
2. Reading Composite+Essay	.39***	<u>.22</u> ***	.51***	.17*
3. Cloze+Essay	.37***	<u>.21</u> ***	.46***	.18**
4. M/C Read+Essay	.38***	<u>.21</u> ***	.51***	.14*
5. Writing Comp+M/C Read	.39***	.12**	.58***	.23***
6. Writing Composite+Cloze	.37***	.14**	.52***	.27***
Read/Write Total	.39***	.14**	.57***	.26***

***p<.0001; **p<.001; *p<.01 (one-tailed).

Note: Underlined correlations are significantly higher than those obtained for Read/Write Total score (p<.05, one-tailed) as evidenced via t-test between correlations for correlated samples Guilford & Fruchter, 1973, p. 167).

*Academy instructor mean rating of four writing abilities; N=504 for batteries with Read/Write Tests only; N=413 for batteries that include Essay Test.

^bAcademy pass/fail index: Graduated (C1,2)=1; Resigned or terminated due to inadequate writing skills or other KSAs (R2,7 & T2)=0; N=423 for batteries with Read/Write Tests only; N=295 for batteries that include Essay Test.

°N=227.

^dFTO mean rating of four writing abilities; N=329 for batteries that include Read/Write Tests only; N=292 for batteries that include Essay Test.

Table 39

Within-Group Validity Evidence for Alternative Reading and Writing Test Batteries

			Alt	ernative	Test Bat	teries		
Criterion Measure subgroup	R/W Total	#1	#2	#3	#4	#5	#6	No. Subjects
Academy Instructor Rating ^a								
Asian	.44	.36	.37	.37	. (.32)	.43	.42	23-28
Black States and States and	.30	.34	.31	.36	.24	.25	.34	53-68
Hispanic	.31	<u>.42</u>	.40	.38	.40	.31	.27	74-86
White	.32	.30	.31	.27	.31	.32	.29	250-303
Male	.43	.40	.41	.36	.40	.43	.38	289-356
Female	.36	.38	.37	.35	.36	.34	.34	123-145
Academy Success/Failure ^b					1			
Black	(.07)	(.20)	(.14)	(.26)	(.02)	(04)	(.19)	35-58
Hispanic	(.16)	.37	.34	.30	.38	(.19)	(.12)	51-64
White	.16	.23	.23	.18	.24	.18	.13	181-258
Male	.13	.20	.20	.20	.19	.12	.14	233-342
Female	(.15)	.32	.29	<u>.29</u>	<u>.30</u>	(.14)	(.15)	62-79
Academy Proficiency Test								
Reack	41	(17)	(24)	(17)	(25)	46	(32)	26
Hispanic	30	(20)	(.27)	(.17)	30	35	(.32)	36
White	.50	(.20)	40	38	38	.55	39	142
Male	. 57	.42	.+0 52	48	51	.72	51	179
Female	.64	.53	.54	.51	.52	.62	.62	48
FIO Kating	(07)	(^0)	(04)	(00)	(00)			20.47
Black	(.07)	(.08)	(.04)	(.06)	(.04)	(.04)	(.09)	39-47
Hispanic	(.17)	(.20)	(.14)	.26	(.05)	(.09)	<u>.2/</u> 22	40-49
White	.25	.19	.17	.15	10	.24	.23	182-205
Male	.29	.20	.21	.21	.18	.26	.29	209-236
Female	.20	(.12)	(.05)	(.07)	(.02)	.18	.24	83-93

Note: All correlations are significant (p<05, one-tailed) except those in parentheses. Underlined coefficients are significantly higher than those obtained for Read/Write Total (p<05, one-tailed).

Alternative Batteries: (1) R/W Total+Essay, (2) Reading Composite+Essay, (3) Cloze+Essay, (4) M/C Read+Essay, (5) Writing Composite+M/C Read, (6) Writing Composite+Cloze.

^aAcademy instructor mean rating of four writing abilities.

^bAcademy pass/fail index: Graduated (C1,2)=1; Resigned or terminated due to inadequate writing skills or other KSAs (R2,7 & T2)=0.

'POST Basic Course Proficiency Test Score.

^dFTO mean rating of four writing abilities.

Test Battery	Asian N=35	Black N=123	Hispanic N=118	White N=471	Female N=232	Male N=552
1. R/W Total+Essay	04	61*	31	+.27	+.01*	+.01*
2. Reading Composite+Essay	10	66*	33	+.30*	07	+.05
3. Cloze+Essay	09	62	27	+.26	05	+.03
4. M/C Read+Essay	14	62	34	+.30	03	+.03
5. Writing Comp+M/C Read	03	55	32	+.26	.00*	+.01
6. Writing Comp+Cloze	.00	55	28	+.24*	01	+.01
Read/Write Total	02	58	32	+.26	03	+.02

 Table 40

 Relative Difficulty of Alternative Reading & Writing Test Batteries

 Mean Z-scores by Race/Ethnicity and Gender

*Significant difference from Read/Write Total mean Z-score, within group repeated measures ANOVA (p<.05, two-tailed).

Note: Based on Sacramento City job applicant sample; Read/Write Test scores standardized by form.

Summary

Four instances of small, statistically significant improvements in validity were detected for the total sample; namely, Batteries #1-#4 in predicting overall academy success/failure. These alternative batteries were also found to improve prediction of this criterion for Hispanics and females, while prediction of Blacks' success/failure was no better or worse. None of the alternative batteries were found to offer reductions in relative difficulty for racial/ethnic minorities, and two batteries (#1 and #2) were found to be more difficult for Blacks than the present battery. Two instances of slight reductions in difficulty were observed for females (Batteries #1 and #5).

Moderator Analysis

Analyses were conducted to examine the extent to which the observed relationships between Reading & Writing Test scores and academy/field training performance were affected by other variables unrelated to reading and writing abilities. More specifically, the analyses were aimed at identifying variables that may have contaminated the criterion measures of performance in basic training and field training, and which may have suppressed or otherwise influenced the obtained validity coefficients.

Several potential moderator variables were examined, as described below. Unfortunately, measures of subjects' prior law enforcement training and experience were not available for this analysis. However, in previous research conducted by POST (Weiner & Berner, 1987) it was found that experience (as measured by total months of experience working at any California law enforcement agency) and training (as measured by hours of POST-certified training) had small correlations with ratings of officers' performance of writing activities, and small or no changes in the magnitudes of Read/Write Test validity coefficients resulted when adjustments for training and experience differences were made.

Time Span Effect

The first potential moderator that was examined in the present study was the amount of *time* between testing and criterion data collection. The rationale here was that over longer periods of time, subjects might have been afforded the opportunity to develop relevant skills, particularly after deciding to pursue a career in law enforcement. Also, those subjects tested years earlier may have been more likely to have prior experience and training. However, time was *not* found to be significantly correlated with academy instructor mean ratings of writing ability, overall academy success/failure, or FTO mean ratings of writing ability; and a very small, but statistically significant correlation was obtained with Proficiency Test scores (r=.03). Therefore, it was concluded that time between testing and criterion data collection was not a moderator of the obtained validities. These results are summarized in Table 41.

		T.	able	41		
	· · ·	Moder	ator	Analysis		
Time	Between	Testing	and	Criterion	Data	Collection

Criterion Measure	n na ngun na sa sa sa kaong na sang na	Time (months) between testing and criterion data collection							
	N	Mean	\$D	Min	Max	r with criterion			
Academy Rating ^a	504	16.10	17.95	0.20	86.86	03			
Academy Pass/Fail ^b	423	22.25	19.20	3.19	86.86	.03			
Academy Proficiency Test	13,347	12.51	7.95	4.01	36.00	.03***			
FTO Rating ^e	292	17.78	15.53	1.28	89.16	.05			

***p<.0001

Note: Academy ending date was used as date of criterion collection for academy performance ratings and overall success/failure; actual dates were available and used for Proficiency Test and FTO rating data collection.

*Academy instructor mean rating on four writing ability scales.

^bAcademy pass/fail index: Completed (C1,2)=1; Failed to complete basic training due to inadequate report writing or other KSAs (R2,7 & T2)=0.

*FTO mean rating on four writing ability scales. Rating date was not available for 37 subjects.

Potential Rating Errors

Several additional potential moderators were examined for FTO ratings only. These involved different potential sources of rater bias, or unfairness, and included: (1) how well the evaluator knew the performance of the officer being rated; (2) ratings of the officers' performance in areas unrelated to reading and writing ability, such as physical fitness/ appearance and personality traits, and (3) characteristics of the evaluator, including race/ethnicity and gender.

Indices of the first two of the above potential moderating factors were correlated with FTO mean ratings of writing ability. Descriptions of these indices and the results are reported in Table 42. The results indicate that there is a positive and significant linear relationship between the magnitude of writing ability ratings given by an evaluator and: (a) the evaluator's familiarity with the ratee's job performance, (b) ratings of the officer's physical fitness and appearance given by the evaluator, and (c) ratings of personality traits of the officer given by the evaluator (r=.34, .34, and .38, respectively). That is, factors that should not be reflective of an officer's writing ability were found to be correlated with FTO ratings of such ability and, thus, were considered to be potential sources of bias.

In light of these findings, analyses were conducted to examine the impact of the above factors upon the earlier obtained correlation between Read/Write total scores and FTO ratings of writing ability. Specifically, FTO ratings were adjusted, or "residualized," by removing that portion of the ratings which is explained by each potential moderating factor and then correlating the residual rating values with test scores. The resulting correlations are called "semipartial correlations" (see Cohen & Cohen, 1975 for further details) and represent estimates of the relationship between test scores and FTO ratings, controlling for significant differences in ratings due to possible rating inaccuracies.

Semipartial correlations between Read/Write scores and adjusted FTO ratings of writing ability are shown in Table 43. As seen in the table, there was no improvement in prediction of FTO ratings when adjustments were made for how well the rater knew the ratee's performance, or for ratings on physical fitness/appearance or personality traits. The original zero-order correlation was .26 (see Table 24), while the semipartials ranged from .23 to .27.

	How well rater knows ratee perf ^a	Physical Fitness/ Appearance ^b	Personality Traits°
N	305	329	329
Mean	2.4	4.0	3.8
SD	0.5	0.6	0.6
Min	2	2.5	2
Max	3	5	5
r with FTO ratings of writing ability	.34	.34	.38

Table 42Moderator AnalysisPotential Sources of Rater Bias in FTO Ratings

Note: All correlations are significant (p<.0001).

Table 43Semipartial CorrelationsRead/Write Total Scores and Adjusted FTO Ratings of Writing Ability

FTO ratings of writing ability adjusted for	N	^r
How well rater knows ratee performance	305	.23
Physical Fitness/Appearance	329	.26
Personality Traits	329	.27
All of the above	305	.24

Note: r = semipartial correlation; all correlations are significant (p<.0001).

^a2=fairly well, 3=very well.

^bMean of ratings on Appearance and Physical Fitness scales

'Mean of ratings on Interpersonal Behavior, Teamwork, and Emotional Self-Control scales.

Rater Characteristics

Another set of analyses was conducted to explore the possibility that characteristics of the FTOs who made the evaluations, i.e., race/ethnicity and gender, somehow influenced or were associated with the field training performance ratings that were given to officers (information regarding the race/ethnicity and gender of academy instructors was not available).

First, a simple comparison was made of the mean ratings of writing ability and overall job performance given by FTOs of different racial/ethnic groups (Asian, Black, Hispanic, White, and other non-White) and gender groups. This simple approach is not suggested as a definitive analysis because the groups are not matched and thus, any significant differences between rater groups may be due to real differences between the rated officers, or they may indeed reflect rater biases. However, a finding of no significant difference between rater groups suggests that, statistical power issues aside, the raters did not favor one group over another by giving higher ratings (although bias could still be present if rater groups gave the same average ratings but the ratees actually differed in performance).

Interestingly, there were *no* statistically significant differences found between mean ratings given by different FTO *racial/ethnic* groups and the magnitudes of the mean ratings were not greatly different. For example, the difference between group mean ratings of writing ability made by Black and White FTOs was only 0.02 scale points; and Hispanic and White FTOs rated officers the same, on average. The difference between Black and White FTO mean ratings of officers' overall job performance was 0.14 scale points, while the difference between Hispanic and White FTO mean ratings of such performance was 0.08 scale points. *Female* FTOs were found to give *significantly higher* ratings than male FTOs on *writing ability*, but not overall job performance.

The mean ratings made by FTOs within each racial/ethnic and gender group are shown in Table 44, along with results of statistical significance tests (one-way ANOVAs). It should be noted that some of the subgroups were small, and thus, statistical power to detect differences was not high.

Rater-by-Ratee Interaction. Unfortunately, there were insufficient numbers of FTOs within each of the racial/ethnic and gender subgroups to facilitate a complete analysis of the interaction between rater and ratee characteristics (i.e., the extent to which FTOs within each subgroup rated members of various subgroups the same or differently). Nevertheless, in order to provide simple descriptive information regarding the rater sample, mean ratings for each rater-ratee race/ethnicity and gender group combination are presented in Tables 45 and 46, respectively.

The sample sizes for White raters and male raters were sufficiently large to facilitate statistical comparisons of ratee group mean ratings within each of these rater groups (one-way ANOVAs). It is noteworthy that *White FTOs did not give significantly different ratings*, on

FTO Rating	Asian	Black	Hispanic	White	Other	Male	Female
Writing Ability N Mean SD	14 3.79 0.63	20 3.81 0.86	42 3.83 0.67	177 3.83 0.58	17 3.68 0.57	239 3.79 0.61	26 4.05 0.69
ANOVA	F=0.24 (4	,265), p=.91	41			F=4.05 (1,26	3), p=.0453*
Overall Performance N Mean SD	13 3.38 0.65	19 3.53 0.77	39 3.59 0.82	172 3.67 0.67	16 3.69 0.60	229 3.63 0.69	26 3.62 0.80
ANOVA	F=0.70 (4	254), p=.59	15	•		F=0.01 (1,25	3), p=.9261

Table 44FTO Mean Ratings of Writing Ability and Overall Performanceby Evaluator Race/Ethnicity and Gender

FTO Ratings	Ra	atee race/ethnic	ity
	Black	Hispanic	White
Writing Ability			
Black N Mean SD	3 2.67 1.53	4 3.88 0.72	12 4.04 0.52
Hispanic N Mean SD	3 3.58 0.72	7 3.79 0.68	28 3.79 0.62
White ^a N Mean SD	24 3.71 0.55	24 3.75 0.60	115 3.88 0.58
Overall Performance			
Black N Mean SD	3 2.67 1.53	4 3.50 0.58	11 3.72 0.47
Hispanic N Mean SD	3 3.67 0.58	5 3.80 0.84	27 3.52 0.85
White ^b N Mean SD	24 3.58 0.58	22 3.64 0.79	113 3.71 0.68

Table 45FTO Mean Ratings of Writing Ability and Overall Performance
by Evaluator and Ratee Race/Ethnicity

^aOne-way ANOVA, white raters only (I.V.=ratee race/ethnicity): F=1.26 (2,162), p=.2855.

^bOne-way ANOVA, white raters only (I.V.=ratee race/ethnicity): F=0.38 (2,156), p=.6828.

	Table 46	
FTO Mean	Ratings of Writing Ability and Overall Performance	Э.
	by Evaluator and Ratee Gender	

FTO Ratings	Ratee gender	
	Female	Male
Writing Ability		
Female N Mean SD	8 4.19 0.46	18 3.99 0.77
Male ^a N Mean SD	66 3.97 0.61	173 3.72 0.60
Overall Performance		
Female N Mean SD	8 3.63 0.52	18 3.61 0.92
Male ^b N Mean SD	61 3.72 0.71	168 3.60 0.68

*One-way ANOVA, male raters only (LV.=ratee gender): F=8.01 (1,237), p=.0050.**

^bOne-way ANOVA, male raters only (I.V.=ratee gender): F=1.51 (1,227), p=.2200.

average, to Black, Hispanic, or White officers. Moreover, the magnitudes of these means were at approximately the same scale level (between 3 and 4). Also interesting is the finding that male FTOs rated female officers significantly higher than they rated male officers in writing ability, but not overall performance. The latter finding suggests that while male FTOs tended to give lower writing ability ratings than female FTOs, on average (as reported above), male and female FTOs were consistent in rating female officers higher in writing ability than male officers. Thus, overall there was no apparent trend for FTOs within racial/ethnic and gender majority groups (Whites and males) to give systematically lower ratings to minority group members.

Summary

Several variables were examined with respect to their effects upon criterion measures of performance in basic training and field training and, in turn, the validity coefficients obtained for Read/Write scores with these criteria. Overall, the results indicated that none of the variables studied were found to moderate the validity results reported earlier. Specifically, no moderating effects were found to be associated with: (a) *time* between testing and criterion data collection, (b) the evaluator's degree of familiarity with the ratee's job performance, (c) the evaluator' ratings of the officer's physical fitness/appearance, or (d) the evaluator's ratings of personality traits of the officer. Additional analyses were conducted to examine the extent to which rater race/ethnicity and gender were associated with the ratings given to officers. Overall, there was no apparent trend for racial/ethnic minorities and females to receive systematically lower ratings than Whites and males when rated by White or male FTOs.

<u>Utility</u>

The above described validity evidence indicates that a statistically significant linear relationship exists between Reading & Writing Test scores and measures of performance, both in the basic academy and in field training. That is, examinees who score lower on the Reading & Writing Test tend to be rated lower in their demonstrated writing abilities, while higher scoring examinees tend to be rated higher on these abilities. A significant relationship was also found between Read/Write scores and overall success or failure in completing basic training, as well as with knowledge of the basic training curriculum as measured by Proficiency Test scores.

While the empirical validity results are important in that they document the jobrelatedness of the test battery, there are additional factors which affect the practical utility of the battery; namely, the base rate for successful job performance (the percentage of employees who would be successful without using the test as a screening device) and the passing rate resulting from the cut score used with the test (selection ratio). This is the classic Taylor-Russell model for assessing the utility of a test (see Cascio, 1982, Ch. 7). Under this model, when validity is held constant, a test will have maximum utility when the base rate is near 50% and when the selection ratio is low. As the selection ratio approaches 100% and/or the base rate departs from 50%, the usefulness of the test decreases until there is little or no gain realized from administering the test.

Thus, it is possible for a highly valid test to have low utility when either the selection ratio or the base rate for successful job performance is high. Likewise, it is possible for a marginally valid test to offer substantial utility when the selection ratio is low (few examinees are selected) and the base rate is near 50%.

Expectancy Tables. Estimates of the utility of Read/Write scores in predicting academy and field training performance were computed for the total sample, as well as by race/ethnicity and gender, in those instances where significant validities were obtained. Five levels of Read/Write Test performance were selected for the analyses representing the lower to upper middle range of test performance in 5-point increments (35-55).³¹ It should be noted that because two of the five participating agencies used the POST Read/Write Battery and the remaining agencies used alternative measures of reading and writing abilities in their operational hiring practices, the *true utility of the tests may be underrepresented* as lower scoring examinees at these agencies were less likely to be hired. Thus, *the results of these analyses should be viewed as gains relative to existing selection practices*.

Utility in Predicting Academy Instructor Ratings

The utility of Reading & Writing Test scores in predicting academy instructor ratings of writing ability for the total validation sample is summarized in Tables 47a and b. Both tables present the percentages of academy students rated as *satisfactory* who scored above and below each of the five Read/Write score levels. In addition, the tables present the percent gain in satisfactory performers, relative to the base rate of academy performance, that would be realized using each of the five Read/Write score levels as a cut score.³²

The difference between the two tables is in how *satisfactory* performance was defined. In the first table, a mean rating corresponding to *adequate* (3) on the 5-point rating scale was used to classify academy students as either minimally acceptable or less than acceptable with regard to writing ability. In the latter table, the median composite rating for the total sample (3.7) was used to classify students as *above average* or *below average* in writing ability. Thus, Table 47a focuses on the utility of Read/Write scores in identifying academy students who demonstrate *minimally acceptable* writing ability, while Table 47b focuses upon *above average* performance. Results pertaining to the former are offered to reflect the minimum

³¹Given that the standard error of measurement for Read/Write total scores is approximately 4 points, it was deemed reasonable to focus on these broad score intervals.

³²Percent gain represents the relative improvement over the base rate of criterion performance and was computed as follows: ((percent adequate or above average/base rate percent)-1)*100.

Table 47Empirical Expectancy TablesReading & Writing Test Scores Predicting Academy Instructor RatingsTotal Sample

Read/Write	% Rated Adea	% Gain	
Cut Score	Achieve cut score	Below cut score	vs. base rate
55	96.5%	88.5%***	4.8%
50	95.5%	86.5%***	3.7%
45	94.9%	84.2%***	3.1%
40	94.0%	79.4%***	2.1%
35	92.9%	77.8%*	0.9%
· · · ·	Base rate = 92	.1% (N=504)	

A. Minimum Acceptable Performance

B. Above Average Performance

Read/Write	% Rated Above 2	% Gain				
Cut Score	Achieve cut score	Below cut score	vs. base rate			
55	∂69.6% ·	40.8%***	29.4%			
50	64.6%	36.3%***	20.2%			
45	62.8%	28.6%***	16.8%			
40	58.9%	20.6%***	9.6%			
35	55.6%	22.2%***	3.3%			
Base rate = 53.8% (N=504)						

Note: Percent gain=((percent rated adequate or above average and achieve cut score/base rate percent)-1)*100. Significant differences (chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score denoted as follows: **p<001; *p<05 (one-tailed).

standards concerns of POST; i.e., screening out candidates who do not possess minimum requisite abilities. Results pertaining to the latter are intended to reflect the concerns of local hiring authorities; i.e., to select the best qualified candidates.

The results in Table 47a are interesting from several standpoints. First, the percentage of students rated as *adequate* who achieved each cut score was found to be significantly higher than the percentage of similarly rated students scoring below each cutoff.³³ That is, Read/Write scores in the range studied were found to distinguish between students performing above and below a minimum acceptable level.

However, because of the high base rate for *adequate* writing ability as measured by academy instructor ratings (92.1%), there was little room for improvement in predicting the criterion defined in this way. Accordingly, the relative percentage gains in academy students identified as *adequate* performers were small, ranging from less than 1% for a score of 35, to a gain of 4.8% for a score of 55. It is noteworthy that the relative gains increased monotonically with test score level, consistent with the observed significant correlation between test scores and academy ratings.

When the focus is shifted from *adequate* performance to *above average* performance, as in Table 47b, substantially larger gains in academy performance are realized at each Read/Write score level. For the total sample, percentage gains in academy students rated as *above average* in writing ability were found to increase monotonically with test score, ranging from 3.3% for a score of 35, to 29.4% for a score of 55.

Within-Group Results. The utility of Read/Write scores in predicting academy ratings within race/ethnicity and gender is summarized in Tables 48a and b in a somewhat different (condensed) format. The first table presents for each group (Blacks, Hispanics, Whites, males and females), the percentage of academy students rated as *adequate* who scored at or above each of the five Read/Write score levels. Notation is made in those instances where such percentages are statistically significantly higher than the percentage of similarly rated students scoring below the cut score. Relative percent gains are also reported in parentheses and were computed as described above. The second table presents the same information for students rated *above average* in writing ability.

The within-group results were somewhat different than those for the total sample, although interpretation of these findings should be tempered by consideration of the relatively small subgroup sample sizes. With regard to the prediction of *adequate* performance in basic training, the results in Table 48a indicate that statistically significant gains in performance were detected for males only; i.e., significantly higher percentages of males who achieved each cut score were rated as *adequate* than those scoring below the cut score. No such significant differences were detected for Blacks, Hispanics, Whites or females at any of the

³³Chi-square analyses were performed in all instances, except where expected cell frequencies were less than five, in which case, Fisher's exact test was performed.

Table 48

Empirical Expectancy Tables Reading & Writing Test Scores Predicting Academy Instructor Ratings by Race/Ethnicity and Gender

	the second s				
R/W Cut Score	Black	Hispanic	White	Male	Female
55	81.8%	95.2%	98.3%	97.6%***	95.1%
	(5.0%)	(6.4%)	(0.6%)	(7.2%)	(-0.1%)
50	86.4%	91.7%	97.8%	95.6% ***	96.4%
	(10.8%)	(2.4%)	(0.1%)	(5.0%)	(1.3%)
45	83.3%	92.0%	98.0%	94.7%***	96.3%
	(6.9%)	(2.8%)	(0.3%)	(4.0%)	(1.1%)
40	82.7%	90.8%	98.2%	93.5%***	96.1%
	(6.1%)	(1.4%)	(0.5%)	(2.7%)	(1.0%)
35	80.3%	89.6%	98.0%	92.3%***	94.9%
	(3.1%)	(0.1%)	(0.3%)	(1.4%)	(-0.3%)
Base rate	77.9%	89.5%	97.7%	91.0%	95.2%
N	68	86	303	356	145

A. Percent Rated Adequate (>=3.0) and Achieve Cut Score (% Gain relative to base rate shown in parentheses)

B. Percent Rated Above Average (>=3.7) and Achieve Cut Score (% Gain relative to base rate shown in parentheses)

R/W Cut Score	Black	Hispanic	White	Male	Female
55	45.5%	66.7%**	70.8%***	66.5%***	80.3%***
	(14.5%)	(68.6%)	(14.7%)	(32.2%)	(26.6%)
50	45.5%	55.6%**	67.7%***	61.8%***	73.8%***
	(14.5%)	(40.5%)	(9.7%)	(22.9%)	(16.3%)
45	47.2%	52.0% **	66.5%***	59.9%***	71.0%***
	(18.9%)	(31.5%)	(7.8%)	(19.2%)	(11.9%)
. 40	44.2%	47.7%**	64.3%***	55.6%***	68.0%**
	(11.4%)	(20.6%)	(4.2%)	(10.5%)	(7.1%)
35	39.3%	44.2%**	62.6%	52.4%***	64.5%
	(-0.9%)	(11.7%)	(1.4%)	(4.2%)	(1.6%)
Base rate	39.7%	39.5%	61.7%	50.3%	63.4%
N	68	86	303	356	145

Note: Percent gain=((percent rated adequate or above average and achieve cut score/base rate percent)-1)*100. Significant differences (chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score (not shown) denoted as follows: ***p<.01; *p<.05 (one-tailed).

five Read/Write score levels, although the observed gains were generally in a positive direction (the nonsignificant findings may be the result of low statistical power due to small sample size).

Substantially larger gains in academy performance are realized within racial/ethnic and gender groups when the focus is shifted to *above average* performance. As seen in Table 48b, significant findings were obtained for Hispanics, Whites, males and females; and positive, but nonsignificant gains in Blacks' performance were yielded (note that the sample size for Blacks was relatively small). These results indicate that significantly higher percentages of *above average* performers scored at or above the cut scores than scored below the cut scores.

The relative percent gains for Hispanics were the most dramatic; such gains were found to increase with test score, ranging from 11.7% for a score of 35, to 66.7% for a score of 55. The relative percent gains obtained for Whites ranged from 1.4% to 14.7%. The percentage gain estimates for Blacks, while encouraging, are not considered reliable in view of the nonsignificant findings.

The pattern of results was comparable for males and females, which is noteworthy in light of the substantially higher base rate for *above average* performance obtained for females (63.4%) versus males (50.3%). The overall trend for each gender group was consistent with the total sample; i.e., academy performance increasing with test score.

Utility in Predicting Academy Success/Failure

Table 49 contains overall academy success rates for academy students scoring above and below each of the five levels.³⁴ Due to the very high base rate for academy success (98.4%) and the modest validity coefficient obtained between test scores and this index of academy performance (r=.14), little or no gain in prediction of this criterion was found across the score levels.

In view of the very high rate of academy success observed in the present study, follow-up analyses of a similar nature were performed on a much larger and broader sample of academy students who were subjects in a previous POST study in which the base rate for successful completion of the basic academy was found to be substantially lower; i.e., 94.7% (see "1987 Study" in the below section entitled "Comparison of Findings with Other Research" for further information). The results of these follow-up analyses indicate that statistically significant gains in student success were realized at each of the five Read/Write score levels, ranging from 1% to 4%. Within-group analyses of these data detected significant gains in academy student success for males and Whites (the largest samples) at all five Read/Write Test score levels, and for Asians, Hispanics and females at certain

 $^{^{34}}$ Results are shown only for the total sample since minority race/ethnicity and gender group validities were not statistically significant (p>.05, one-tailed).

Table 49

Empirical Expectancy Table Reading & Writing Test Scores Predicting Academy Success/Failure Total Sample

Read/Write	% Succ	% Gain				
Cut Score	Achieve cut score	Below cut score	vs. base rate			
55	100%	97.0%*	1.7%			
50	99.6%	96.4%*	1.3%			
45	99.0%	96. 3%	0.7%			
40	98.7%	96.2%	0.3%			
35	98.5%	94.1%	0.2%			
Base rate = 98.4% (N=423)						

Note: Percent gain=((percent successful and achieve cut score/base rate percent)-1)*100. Significant differences (Fisher's exact test) between percent achieve cut score vs. percent below cut score (not shown) denoted as follows: *p<.05 (one-tailed).

*Academy success/failure index: Graduated (C1,2)=1; Resigned or terminated due to inadequate writing skills or other KSAs (R2,7 and T2)=0.

Read/Write score levels. Positive but nonsignificant gains in student success were also observed for blacks. These results are shown in Appendix H, Tables H-1 and H-2.

Utility in Predicting Academy Proficiency Test Scores

The utility of Reading & Writing Test scores in predicting performance on the Proficiency Test is summarized in Tables 50a and b. In Table 50a, *adequate* performance was defined as a Proficiency Test score greater than, or equal to the 25th percentile (44.1) or higher.³⁵ In Table 50b, *above average* performance was defined as a Proficiency Test score greater than, or equal to the median (52.4).

With regard to the prediction of *adequate* performance, Read/Write total scores were found to produce monotonic gains in Proficiency Test performance ranging from 1.8% to 16.8%. Gains in *above average* performance on the Proficiency Test ranged from 2.8% to 37.1%

Within-Group Results. Significant gains in *adequate* performance on the Proficiency Test were realized for all racial/ethnic and gender subgroups studied; i.e., American Indians, Asians, Blacks, Filipinos, Hispanics, Whites, males and females. Interestingly, the largest such gains were found for Blacks, ranging from 4% to 42.8%. Within-group gains in *above average* Proficiency Test performance were even higher for all subgroups, ranging to over 60% for Blacks, over 50% for Asians and Hispanics, over 40% for Filipinos and females, and over 30% for American Indians and males. These results are shown in Tables 51a and b.

Utility in Predicting FTO Ratings

Tables 52a and b contain results regarding performance in *field training*. Table 52a shows the percentage of examinees scoring above and below each of the five test score levels who were subsequently rated by their FTOs as *adequate* in demonstrated writing ability.³⁶ As seen in this table, there is a very high base rate for *adequate* performance for the total sample (93.9%). As a result, the percentage gains in officers rated as *adequate* in writing ability were found to be small, ranging from less than 1% to approximately 3%. Statistically significant gains in performance were detected only in the 45-55 Read/Write Test score range.

Again, when the focus is shifted from *adequate* performance to *above average* performance, the gains in field training performance associated with Read/Write scores are greater. The results in Table 52b indicate that for the total sample, such gains increase monotonically and range from 1% for a score of 35, to 24.8% for a score of 55. Also as above, gains in performance were statistically significant only in the 45-55 score range.

³⁵This is consistent with the minimum passing score that was established for the POST Basic Course Waiver Examination.

³⁶Within-race/ethnicity results are not presented since minority group validities were not statistically significant.

Table 50Empirical Expectancy TablesReading & Writing Test Scores Predicting Academy Proficiency Test ScoresTotal Sample

Read/Write Cut Score	% Scoring at or ab on Proficiency	ove 25th Percentile Test (>=44.1)	% Gain vs. base rate			
	Achieve R/W score					
55	89.9%	66.4%***	16.8%			
50	85.9%	60.4%***	11.6%			
45	82.3%	55.1%***	7.0%			
40	79.8%	49.0%***	3.8%			
35	78.3%	44.1%***	1.8%			
Base rate = 76.9% (N=13,347)						

A. Minimum Acceptable Performance

B. Above Average Performance

Read/Write Cut Score	% Scoring at o on Proficienc	% Gain vs. base rate				
	Achieve R/W Below R/W score					
55	69.7%	35.5%***	37.1%			
50	63.0%	28.4%***	23.9%			
45	57.6%	23.1%***	13.3%			
40	54.1%	19.9%***	6.4%			
35	52.2%	17.5%***	2.8%			
Base rate = 50.8% (N=13,347)						

Note: Percent gain=((percent rated adequate or above average and achieve cut score/base rate percent)-1)*100. Significant differences (chi-square) between percent achieve cut score vs. percent below cut score denoted as follows: ***p<001 (one-tailed).

Table 51

Empirical Expectancy Tables Reading & Writing Test Scores Predicting Proficiency Test Scores by Race/Ethnicity and Gender

A. Percent Scoring at/above 25th percentile on Proficiency Test (>=44.1) a	nd Achieve R/W	Cut Score
(% Gain relative to base rate shown in parentheses)	

R/W Cut Score	Am. Indian	Asian	Black	Filipino	Hispanic	White	Male	Female
55	94.8%***	91.1%***	82.4%***	87.7%***	83.7%***	90.9%***	91.1%***	83.5%***
	(19.2%)	(24.0%)	(42.8%)	(31.2%)	(26.1%)	(11.8%)	(16.1%)	(22.7%)
50	94.8%***	87.3%***	72.3%***	83.7%***	79.2%***	87.7%***	87.3%***	78.2%***
	(19.2%)	(18.9%)	(25.3%)	(25.1%)	(19.3%)	(7.8%)	(11.2%)	(15.0%)
45	88.4%***	82.1%***	67.0%***	75.0%***	74.1%***	84.9%***	83.8%***	73.6%***
	(11.1%)	(11.8%)	(16.1%)	(12.1%)	(11.7%)	(4.4%)	(6.7%)	(8.2%)
40	85.2%***	77.6%***	63.2%***	69.9%*	70.7%***	83.2%***	81.5%***	70.6%***
	(7.1%)	(5.6%)	(9.5%)	(4.5%)	(6.6%)	(2.3%)	(3.8%)	(3.7%)
35	82.0%***	75.1%**	60.0%***	68.7%*	68.5%***	82.2%***	80.0%***	68.8%***
	(3.0%)	(2.2%)	(4.0%)	(2.7%)	(3.2%)	(1.1%)	(1.9%)	(1.1%)
Base rate	79.6%	73.5%	57.7%	66.9%	66.4%	81.3%	78.5%	68.0%
N	137	407	996	154	1,821	9,495	11,149	1,882

B. Percent Scoring at/above Median on Proficiency Test (>=52.4) and Achieve R/W Cut Score (% Gain relative to base rate shown in parentheses)

R/W Cut Score	Am. Indian	Asian	Black	Filipino	Hispanic	White	Male	Female
55	75.9%***	67.9%***	46.6%***	52.6%***	56.7%***	72.6%***	72.5%***	55.1%***
	(36.8%)	(51.7%)	(62.8%)	(47.4%)	(56.1%)	(28.5%)	(37.0%)	(41.6%)
50	77.9%***	61.6%***	41.2%***	47.8%***	49.6%***	66.4%***	65.7%***	48.4%***
	(40.5%)	(37.8%)	(44.0%)	(33.9%)	(36.5%)	(17.5%)	(24.0%)	(24.4%)
45	68.4%***	54.6%***	35.9%***	41.4%**	44.4%***	61.7%***	60.1%***	43.3%***
	(23.3%)	(22.2%)	(25.6%)	(15.9%)	(22.3%)	(9.3%)	(13.5%)	(11.4%)
40	60.7%***	49.1%***	32.4%***	38.3%*	40.4%***	58.9%***	56.4%***	41.0%***
	(9.3%)	(9.9%)	(13.2%)	(7.4%)	(11.1%)	(4.2%)	(6.5%)	(5.4%)
35	57.1%*	47.0%***	30.4%***	36.1%	37.9%***	57.5%***	54.5%***	39.4%***
	(3.0%)	(4.1%)	(6.1%)	(1.0%)	(4.2%)	(1.8%)	(2.9%)	(1.3%)
Base rate	55.5%	44.7%	28.6%	35.7%	36.4%	56.5%	52.9%	38.9%
N	137	407	996	154	1,821	9,495	11,149	1,882

Note: Percent gain=((percent rated adequate or above average and achieve cut score/base rate percent)-1)*100. Significant differences (chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score (not shown) denoted as follows: ***p<.001; **p<.01; *p<.05 (one-tailed).

Table 52Empirical Expectancy TablesReading & Writing Test Scores Predicting FTO RatingsTotal Sample

Read/Write Cut Score	% Rated Adequate (>=3.0)		% Gain	
	Achieve cut score	Below cut score	vs. base rate	
55	96.6%	91.8%*	2.9%	
50	96.1%	90.2%*	2.3%	
45	96.0%	87.7%**	2.2%	
40	94.4%	90.2%	0.6%	
35	94.0%	92.3%	0.1%	
Base rate = 93.9% (N= 329)				

A. Minimum Acceptable Performance

B. Above Average Performance

Read/Write Cut Score	% Rated Above Average (>=3.8)		% Gain	
	Achieve cut score	Below cut score	vs. base rate	
5,5	62.6%	40.1%***	24.8%	
50	55.8%	40.7%**	11.3%	
45	52.8%	42.0%*	5.3%	
40	51.4%	41.5%	2.5%	
35	50.6%	38.5%	1.0%	
Base rate = 50.2% (N= 329)				

Note: Percent gain=((percent rated adequate or above average and achieve cut score/base rate percent)-1)*100. Significant differences (chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score denoted as follows: ***p<.001; *p<.05 (one-tailed).

Within-Sex Results. Tables 53a and b present within-sex results with respect to *adequate* and *above average* performance in field training, respectively. Only one statistically significant finding was obtained for males (at the 45 score level) and none were obtained for females in predicting *adequate* field training performance. This result is not surprising given the high base rate for this criterion. Only males' *above average* field training performance was predicted, and only in the 45-55 test score range; males' percentage gains in *above average* performers ranged from less than 1% for a score of 35, to 34.5% for a score of 55.

Summary

Read/Write scores were found to offer utility in identifying *above average* performers in both basic training and field training, and were found to be useful to a lesser extent in improving upon the selection of *adequate* performers. For example, when a cut score of 45 on the Read/Write Test was applied to the validation sample, the percentage of academy students rated as *above average* in writing ability increased by 17%; the percentage of students scoring *above average* on the Proficiency Test increased by 13%; and the percentage of field trainees rated as *above average* in writing ability increased by 5%. When the same cut score was applied to predict *adequate* performance on these criteria, the corresponding performance gains were 3%, 7%, and 2%, respectively.

Within-group results indicate that Read/Write scores have utility in improving the selection of *above average* performers in *basic training* for all racial/ethnic and gender subgroups studied for at least one criterion measure. The within-group results for field training performance were inconclusive due to small sample sizes.

Interpretation of the above utility results should be tempered by consideration of the high base rates for *adequate* performance in the academy and in field training. That is, the relatively high performance base rates left little room for improvement and, thus, only small gains in utility were possible. Also, as stated earlier, the above gains are relative to existing personnel selection procedures.

Furthermore, limitations of the obtained criterion measures of academy and job performance should also be considered. For example, while the academy instructor and FTO rating scales were designed to capture important aspects of students' and officers' report writing skills, they are not direct assessments of writing proficiency and are limited by: (a) academy instructors' and FTOs' expertise in judging the quality of writing, and (b) differences in the types and amounts of writing required by various academies and on the job. A direct assessment of students' and officers' writing proficiency obtained under standard and realistic conditions, and evaluated by experts would be expected to provide a more accurate criterion measure of their writing competencies and, in turn, more accurately portray the utility of Reading & Writing Test scores.

Table 53

Empirical Expectancy Tables Reading & Writing Test Scores Predicting FTO Ratings by Gender

R/W Cut Score	Males	Females
55	95.4% (2.8%)	100% (3.3%)
50	94.7% (2.1%)	100% (3.3%)
45	94.9% * (2.3%)	98.6% (1.9%)
40	93.7% (1.0%)	96.3% (-0.4%)
35	92.9% (0.1%)	a
Base rate	92.8%	96.8%
N	236	93

A. Percent Rated Adequate (>=3.0) and Achieve Cut Score (% Gain relative to base rate shown in parentheses)

B. Percent Rated Above Average (>=3.8) and Achieve Cut Score (% Gain relative to base rate shown in parentheses)

R/W Cut Score	Males	Females
55	58.7%*** (34.5%)	73.7% (10.5%)
50	51.7%*** (18.4%)	67.3% (0.9%)
45	47.5%* (8.7%)	66.2% (-0.7%)
40	45.6% (4.6%)	65.9% (-1.2%)
35	43.8% (0.4%)	a
Base rate	43.6%	66.7%
N	236	93

Note: Percent gain=((percent rated adequate or above average and achieve cut score/base rate percent)-1)*100. Significant differences (chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score (not shown) denoted as follows: ***p<.001; *p<.05 (one-tailed).

^aNot computed due to small number of examinees scoring below this level (N=3).

89

Comparison of Findings with Other Research

The overall Reading & Writing Test validity results obtained in the present study, i.e., that test scores are predictive of performance in basic training and on the job, are consistent with the results of research conducted by POST over the last decade, as well as with cumulative validity evidence for tests of verbal ability in predicting performance in law enforcement occupations.

Other POST Studies

As indicated earlier, several empirical studies of the Reading & Writing Test have been conducted by POST. These studies are briefly outlined below.

1981 Study: In the original Reading & Writing Test validation research conducted by POST (Honey & Kohls, 1981), Reading & Writing Tests and an Essay Test were administered to approximately 300 basic academy students at five academies near the beginning of training, and measures of their performance in the academy were obtained later, near the end of training. The academy performance measures were two written tests of knowledge of the basic training curriculum; namely, the POST Basic Academy Proficiency Test and a locally developed test, specific to each academy. It should be noted that the test battery did *not* include a traditional multiple-choice Reading Comprehension Test at this time, and the Essay test was scored using an *analytical* procedure rather than the current *holistic* scoring approach. Essay Test score predictions of academy test performance were further analyzed after the 1981 validation report was published and the resulting validities are reported in the below summary table (Table 54).

1983 Study: A second academy study was conducted which served as the basis for the development of a test user's manual in 1983. By this time, a multiple-choice Reading Comprehension Test had been developed and added to the battery. The study entailed administering the Reading & Writing Test to 480 academy students at 10 academies, and then collecting measures of their performance in the academy (the academy performance measures were the same achievement tests used in the 1981 study).

1987 Study: A predictive criterion-related validation study was conducted in 1986-87 as a follow-up evaluation of the operational Reading & Writing Test battery (Weiner & Berner, 1987). Reading & Writing Test scores were retrieved from POST's operational test program files for 1270 examinees and measures of their subsequent performance in basic training and on-the-job were collected. The Essay Test was not included in this study. Academy performance was measured by scores on the POST Basic Academy Proficiency Test. Performance on the job was evaluated using several specially developed measures, including: (a) field training officer (FTO) ratings of officer/trainees' performance of job duties that involve writing; Behaviorally Anchored Rating Scales (BARS) were used to evaluate officer/trainees; (b) overall success or failure in completing field training (a dichotomous pass/fail variable); (c) patrol supervisor ratings of tenured officers' performance of job duties that involve writing (again using BARS); and (d) overall success/failure in completing probation.

Supplemental analyses of Reading & Writing Test score predictions of overall success or failure in completing basic training (graduated vs. failed/withdrew for academic reasons) were conducted after the 1987 validation report was published and the results are included in the below summary tables (Tables 54-56). The success criterion data were collected by POST from 27 basic academies between 1986 and 1987. These academies indicated on a special data collection form whether each student graduated, withdrew for academic reasons, withdrew for other reasons, failed for academic reasons, failed for other reasons, or recycled to the next academy. Reading & Writing Test scores were retrieved from POST's operational testing program files for students who had previously taken the test, resulting in a validation analysis file comprised of 1271 students for whom test scores and academy success/failure data were available.

Current Study: This refers to the results of the present research described earlier in this report, wherein Reading & Writing Test and Essay Test scores were examined with respect to their predictions of subsequent performance in basic training (academy instructor ratings of writing ability, overall success/failure in completing basic training, and POST Proficiency Test scores) and in field training (FTO ratings of writing ability and overall success/failure in completing field training).

The results of these studies are described below and are summarized in the following tables with regard to overall validity evidence, within-group validity, and differential prediction analysis results, respectively.

Overall Validity Evidence. As seen in Table 54, Read/Write total and subtest scores and Essay Test scores were found to predict *academy performance* in each of the studies and with each measure of academy performance, with one exception (Clarity predicting overall academy success/failure in the current study). The smaller validities obtained in the present study for Read/Write scores in predicting overall academy success/failure may be due to the smaller and more restricted sample that was obtained (N=423 officers, 5 agencies) versus the 1987 study (N=1271 officers and over 50 agencies).
Table 54 Summary of Validity Evidence for POST Reading & Writing Tests

Criterion Measure	R/W Total	Read STD	Write STD	Clarity	Spell	Vocab	M/C Read	Cloze	Essay
Academy Performance									
POST Proficiency Test 1981 study ^a 1983 study ^b 1987 study ^c Current study ^d	.52*** .60*** .56*** .47***	N/A .66*** .56*** .49***	.40*** .41*** .42*** .34***	.28*** .38*** .35*** .25***	.21*** .09* .20*** .14***	.41*** .46*** .39*** .36***	N/A .61*** .52*** .45***	.50*** .54*** .44*** .40***	.18* N/A N/A .28***
Academy-Specific Test 1981 study [*] 1983 study ^b	.53*** .54***	N/A .55	.45*** .40***	.33*** .33***	.25*** .19***	.41*** .36***	N/A .51***	.47*** .44***	.25*** N/A
Instructor Ratings Current study ^e	.39***	.36***	.35***	.24***	.28***	.26***	.33***	.31***	.31***
Academy success/failure 1987 study ^f Current study ⁸	.21*** .14**	.24*** .12**	.14*** .12**	.09*** (.04)	.09*** .13**	.13*** .10*	.20*** .09*	.21*** .13**	N/A .21**
Job Performance									
FTO ratings 1987 study ^h Current study ⁱ	.38*** .26***	.27** .21***	.40*** .24***	.24** .13**	.36*** .26***	.24** .14**	.27** .15**	.18* .23***	N/A (.09)
Patrol Supervisor ratings 1987 study ^j	.22***	.19***	.20***	.18***	.15**	.12**	.17***	.14**	N/A
Field training success/failure 1987 study ^k Current study ^l	.10*** (.05)	.12*** (.06)	.07* (.04)	.10*** (.08)	(.03) (01)	(.02) (.02)	.09** (.07)	.11*** (.03)	N/A (.05)
Probation success/failure 1987 study ^m	.13***	.13***	.10**	.09**	.07*	(.05)	.10***	.13***	N/A

***p<.001; **p<.01; *p<.05 (one-tailed).

^aRead/Write Test N=218-320; Essay N=147-149. Reading test includes two cloze tests (no multiple-choice test). Essay test was scored using an analytical method.

^bN=480.

^cN=1270.

^dRead/Write Test N=13,347; includes all available Read/Write scores matched to Proficiency Test scores obtained between Aug83-Feb92. Essay Test N=227; data collected for current study.

^eRead/Write Test N=504; Essay N=413. Criterion is FTO mean rating on 4 writing abilities. Essay scored using holistic procedure.

^fSupplemental study; N=1271. Pass/fail index: Graduated=1; failed/withdrew for academic reasons=0. Correlations are point-biserials.

^gRead/Write Test N=423; Essay N=295. Pass/Fail index: Completed (C1,2)=1; resigned or terminated due to inadequate writing skills or other KSAs (R2,7 & T2)=0. Correlations are point-biserials.

^hN=103. Criterion is FTO rating of officer performance of job duties that involve writing.

ⁱRead/Write Test N=329; Essay N=292. Criterion is FTO mean rating on 4 writing abilities.

N=382. Criterion is patrol supervisor rating of tenured officer performance of job duties that involve writing.

^kN=1062. Pass/fail index: Completed (C1-3)=1; Failed to complete (R2,3 or T2,3 or F2,3)=0. Correlations are point-biserials.

¹Read/Write Test N=403; Essay N=336. Pass/fail index: Completed (C1-3)=1; Resigned or terminated due to inadequate report writing skills, analytical skills, or other KSAs (R3,4,5,7 or T3,7)=0. Correlations are point-biserials.

^mN=895. Pass/fail index: Completed (C1-3)=1; Failed to complete (R2,3 or T2,3 or F2,3)=0. Correlations are point-biserials.

With regard to *performance on-the-job*, Read/Write total and subtest scores were found to predict each of the different indices of such performance in five of six analyses. That is, with few exceptions, FTO ratings of trainees' writing ability, patrol supervisor ratings of tenured officers' performance of writing-related job duties, and officers' overall success/failure in completing both field training and probation were predicted (Spelling and Vocabulary were not found to predict overall field training success/failure; Vocabulary was not predictive of probation success/failure). Again, the instance of nonsignificant prediction of overall field training success/failure in the current study may be due to the smaller and more restricted sample (N=403 officers, 5 agencies) versus the 1987 study (N=1062 officers and over 50 agencies).

Essay Test score relationships with job performance were not examined in the research conducted prior to the current study. As indicated earlier in the results of the current study, Essay Test scores were found to predict certain FTO ratings (Organization & Narrative and Mechanics), and were *not* found to predict overall field training success/failure.

Within-Group Validity. Validity results within racial/ethnic and gender groups obtained for Read/Write total scores in the above described studies are summarized in Table 55. These results indicate that Read/Write total scores were found to predict *academy performance* as measured by *achievement test scores* and *instructor ratings* of writing ability for all racial/ethnic and gender subgroups included in each of the studies. Overall *academy success/failure* was predicted for Asians, Hispanics, Whites, males and females, and positive, but nonsignificant correlations were obtained for Blacks and Filipinos in the 1987 study. Again, the nonsignificant validities obtained in the present study in predicting overall academy success/failure are likely due to the smaller and more restricted sample that was obtained.

The validity evidence for the prediction of *job performance* within racial/ethnic and gender subgroups is generally inconclusive due to the small samples of minority group members that were available for the research studies (i.e., statistical power was generally low). Positive correlations were obtained for Blacks, Hispanics, Whites, females, and males between Read/Write scores and each of the job performance measures (FTO ratings, patrol supervisor ratings, field training success/failure, and probation success/failure); and positive correlations were obtained for Asians with FTO ratings and overall field training success/failure. However, the obtained correlations were statistically significant only for the larger samples (males and Whites), with few exceptions; i.e., females' FTO ratings and overall field training success/failure were significantly predicted; Hispanics' patrol supervisor ratings were predicted; and Blacks' field training success/failure was predicted. In view of these results, further study of within-group relationships between test scores and job performance is warranted.

Table 55Summary of Within-Group Validity for the
POST Reading & Writing Test Battery

Criterion Measure	Am. Indian	, Asian	Black	Filipino	Hispanic	White	Male	Female
Academy Performance				-			•	
POST Proficiency Test 1981 study ^a 1983 study ^b 1987 study ^e Current study ^d	.57***	.73** .68** .52***	.62** .55*** .53*** .40***		.49** .66*** .49*** .40***	.46*** .57*** .54*** .45***	.53*** .63*** .55*** .48***	.62*** .61*** .66*** .49***
Academy-Specific Test 1981 study ^e 1983 study ^f	•	.51* .66**	.57** .35*		.47** .56***	.51*** .51***	.57*** .54***	.56*** .63***
Instructor Ratings Current study ^g		.44*	.30**		.31**	.32***	.43***	.36***
Overall success/failure 1987 study ^h Current study ⁱ		.38** (.00)	(.13) (.07)	(.15)	.22** (.16)	.16*** .16**	.24*** .13**	.16* (.15)
Job Performance							•	
FTO ratings 1987 study ⁱ Current study ^k		(.12)	(.07)		(.17)	.30** .25***	.38*** .29***	.20*
Patrol Supervisor ratings 1987 study ¹			(.31)		.38*	.19**	.22***	(.23)
Field training success/failure 1987 study ^m Current study ⁿ		(.05)	.20* (.15)		(.02) (.04)	.07* (.05)	.09** (.06)	.17* (.05)
Probation success/failure 1987 study°		•	(.15)		(.04)	.09**	.12***	(.15)

***p<.001; **p<.01; *p<.05 (one-tailed); Correlations are not reported in instances where N was very small (<10).

*Asian N=12, Black N=19, Hispanic N=34, White N=154, Male N=191, Female N=28. Reading test includes two cloze tests (no multiple-choice test).

^bAsian N=17, Black N=41, Hispanic N=60, White N=346, Male N=405, Female N=75.

Black N=111, Hispanic N=137, White N=953, Male N=1103, Female N=167.

^dAmerican Indian N=137, Asian N=407, Black N=996, Filipino N=154, Hispanic N=1821, White N=9495, Male N=11149, Female N=1882. Includes all available Read/Write scores matched to Proficiency Test scores obtained between Aug83-Feb92.

*Asian N=12, Black N=19, Hispanic N=35, White N=152, Male N=190, Female N=28.

^fAsian N=17, Black N=41, Hispanic N=60, White N=346, Male N=405, Female N=75.

⁸Asian N=28, Black N=68, Hispanic N=86, White N=303, Male N=356, Female N=145. Criterion is FTO mean rating on 4 writing abilities.

^hAsian N=48, Black N=107, Filipino N=18, Hispanic N=181, White N=886, Male N=1058, Female N=194. Pass/fail index: Graduated=1; failed/withdrew for academic reasons=0. Correlations are point-biserials.

ⁱAsian N=27, Black N=58, Hispanic N=64, White N=258, Male N=342, Female N=79. Pass/Fail index: Completed (C1,2)=1; resigned or terminated due to inadequate writing skills or other KSAs (R2,7 & T2)=0. Correlations are point-biserials.

^jWhite N=81, Male N=86. Criterion is FTO rating of officer performance of job duties that involve writing.

*Asian N=21, Black N=47, Hispanic N=49, White N=205, Male N=236, Female N=93. Criterion is FTO mean rating on 4 writing abilities.

Black N=15, Hispanic N=31, White N=311, Male N=337, Female N=37. Criterion is patrol supervisor rating of tenured officer performance of job duties that involve writing.

^mBlack N=95, Hispanic N=118, White N=795, Male N=929, Female N=133. Pass/fail index: Completed (C1-3)=1; Failed to complete (R2,3 or T2,3 or F2,3)=0. Correlations are point-biserials.

^aAsian N=24, Black N=64, Hispanic N=64, White N=236, Male N=275, Female N=126. Pass/fail index: Completed (C1-3)=1; Resigned or terminated due to inadequate writing, analytical, or other KSAs, or performance level unknown (R3,4,5,7 or T3,7)=0. Correlations are point-biserials.

^oBlack N=89, Hispanic N=108, White N=651, Male N=779, Female N=116. Pass/fail index: Completed (C1-3)=1; Failed to complete (R2,3 or T2,3)=0. Correlations are point-biserials.

Differential Prediction. Studies of the differences between the prediction equations yielded by Read/Write Test scores for racial/ethnic minorities versus Whites have consistently found that Read/Write Test scores are *not unfair* to minorities in predicting their performance in either *basic training* or *field training*. That is, despite the fact that significantly different prediction equations were yielded for racial/ethnic minorities, their predicted performance was found to be either consistent with actual performance, or was *over*predicted. These results were replicated several times for Blacks and Hispanics, and less frequently for Asians. See Table 56.

The results across studies of male-female differences in test score predictions indicate that females' *academy performance* was *over*predicted when performance was measured by achievement test score or overall success/failure in completing basic training; but their performance was *under*predicted when performance was measured by instructor ratings of writing ability. Interestingly, females' *job performance* was also underpredicted when measured by FTO ratings of writing ability (using the same scales as in the academy instructor booklet), but their performance was *not under*predicted when measured by patrol supervisor ratings of performance of job duties that involve writing (these were different scales than above; i.e., BARS), field training success/failure, or probation success/failure.

Table 56Summary of Differential Prediction Analyses of the
POST Reading & Writing Test Battery

Criterion Measure	Am. Indian	Asian	Black	Filipino	Hispanic	Female
Academy Performance						
POST Proficiency Test 1981 study ^a 1983 study ^b 1987 study ^c Current study ^d	· · · S	Int (+) SE (+)	Int (+) Int (+) Int (+) SE (+)	Int (+)	SE (+) Int (+) Int (+) SE (+)	Int (+) Int (+) S (+) SE (+)
Academy-Specific Test 1981 study ^e 1983 study ^f		None	Int (+) SE (+)		Int (+) SE	Int (+) Int (+)
Instructor Ratings Current study ⁸		None	SE (+)		SE	Int (-)
Overall success/failure 1987 study ^h Current study ⁱ		SE SE	SE SE	SE ·	SE SE	SE (+) SE
Job Performance						
FTO ratings 1987 study ^j Current study ^k		SE	SE	•	None	Int (-)
Patrol Supervisor ratings 1987 study ¹	•		None	•	Int (+)	None
Field training success/failure 1987 study ^m Current study ⁿ	•		S (+)		Int	None
Probation success/failure 1987 study°		•	Int (+)		None	None

Note: Significant differences between minority and majority group regression parameters are denoted as follows: SE=standard errors, S=slopes, Int=intercepts, "None" denotes no parameter differences; "+" denotes overprediction and "-" denotes underprediction of minority group performance. Residual analyses in original 1987 study were conducted relative to majority group (rather than total sample). Analyses were not conducted in instances where N was very small (<10).

*Black N=19, Hispanic N=34, Female N=28 (White N=154, Male N=191). Reading test includes two cloze tests (no multiple-choice test).

^bAsian N=17, Black N=41, Hispanic N=60, Female N=75 (White N=346, Male N=405).

"Black N=111, Hispanic N=137, Female N=167 (White N=953, Male N=1103).

^dAmerican Indian N=137, Asian N=407, Black N=996, Filipino N=154, Hispanic N=1821, Female N=1882 (White N=9495, Male N=11149).

Black N=19, Hispanic N=35, Female N=28 (White N=152, Male N=190).

^fAsian N=17, Black N=41, Hispanic N=60, Female N=75 (White N=346, Male N=405).

⁸Asian N=28, Black N=68, Hispanic N=86, Female N=145 (White N=303, Male N=356). Criterion is FTO mean rating on 4 writing abilities.

^hAsian N=48, Black N=107, Filipino N=18, Hispanic N=181, Female N=194 (White N=886, Male N=1058). Pass/fail index: Graduated=1; failed/withdrew for academic reasons=0.

Asian N=27, Black N=58, Hispanic N=64, Female N=79 (White N=258, Male N=342). Pass/Fail index: Completed (C1,2)=1; resigned or terminated due to inadequate writing skills or other KSAs (R2,7 & T2)=0.

There were insufficient numbers of minorities to complete an analysis with this criterion.

^kAsian N=21, Black N=47, Hispanic N=49, Female N=93 (White N=205, Male N=236). Criterion is FTO mean rating on 4 writing abilities.

¹Black N=15, Hispanic N=31, Female N=37 (White N=311, Male N=337). Criterion is patrol supervisor rating of tenured officer performance of job duties that involve writing.

"Black N=95, Hispanic N=118, Female N=133 (White N=795, Male N=929). Pass/fail index: Completed (C1-3)=1; Failed to complete (R2,3 or T2,3 or F2,3)=0.

"Analysis was not conducted due to nonsignificant validities obtained for total sample and all sub-groups.

"Black N=89, Hispanic N=108, Female N=116 (White N=651, Male N=779). Pass/fail index: Completed (C1-3)=1; Failed to complete (R2,3 or T2,3 or F2,3)=0.

Cumulative Research

The importance of general verbal ability for the successful performance of patrol officer work in California, as well as in other states across the nation, is well documented. In the *California Entry-level Law Enforcement Officer Job Analysis* (Kohls, Berner & Luke, 1979), over two dozen important reading- and writing-oriented job tasks were identified which are performed by a majority of officers statewide, and which served as a basis for establishing the importance of both reading and writing abilities. Subsequent statewide surveys and analyses conducted by POST as part of the 1981 study identified specific writing demands and the types and level of materials commonly read by patrol officers in the state. Examples of the importance of reading and writing abilities for police work in other states are given in Gael's *Job Analysis Handbook* (Bernardin, 1988, Chapter 10.8), where results are summarized for police officer job analyses conducted in 10 different jurisdictions. These job analyses identified report writing activities and verbal ability as important components of the job, and the author concludes that, despite subjective differences in job analytic methodologies, there is a high degree of commonality in police work across jurisdictions.

The empirical validity of verbal ability tests in predicting performance in law enforcement occupations has been summarized by Hirsh, Northrop & Schmidt (1986). In their validity generalization study, analyses were conducted in which cognitive test validity results were aggregated across a number of studies, primarily for occupational group 375 in the Dictionary of Occupational Titles -- Police and Detectives in Public Service (U.S. Dept of Labor, 1977). Results of their research for *verbal ability* tests are summarized below. According to the test classification scheme used in the study, *verbal ability* tests include traditional Reading Comprehension, Vocabulary, Grammar, Spelling, Word Fluency, and Sentence Completion.

26 validity coefficients were aggregated for verbal ability tests as predictors of performance in *training* (N=3,943); the mean validity coefficient was .369; the estimated true validity of such tests was .62 to .64 (corrected for range restriction and unreliability in both the criterion and predictor).

18 validity coefficients were compiled for verbal ability tests in predicting *job* proficiency (N=2,207), which in virtually all cases was measured by supervisory ratings of performance; the mean validity coefficient was .089; the estimated true validity was .18 to .22.

On the basis of their analyses, the authors concluded that the validity of verbal ability tests is generalizable across law enforcement jobs in this category and that cognitive ability tests are excellent predictors of performance in *job training* programs. They further reasoned that the relatively lower validities obtained in predicting *job proficiency* may be due to problems with the criteria; i.e., the difficulty of obtaining reliable and valid measures of job performance for patrol officers, particularly in view of the often unobserved/unsupervised performance of their duties. The fact that higher validities were obtained in the current POST

97

study in predicting on-the-job performance may be reflective of the extensive steps that were taken to collect these data; e.g., specially developed rating instruments with behavioral/job-related rating dimensions, rater training, carefully controlled data collection, etc.

Overall, the results of the Reading & Writing Test validation research conducted by POST are consistent with the above described validity generalization research in that (a) significant prediction of performance in both training and on the job was found, (b) validities were within a comparable range of magnitude, and (c) training performance was better predicted than job performance.

SUMMARY AND CONCLUSIONS

The study addressed major issues pertaining to the empirical validity of scores on the POST Reading & Writing Test Battery, an Essay Test, and several alternative configurations of these tests as predictors of subsequent performance in basic training and on the job (in field training). In addition, test-criterion relationships were examined within racial/ethnic and gender groups, and several potential sources of contamination in the criterion measures were examined with respect to their potential moderating effects upon these relationships. The practical utility of Reading & Writing Test scores was also examined in terms of expected gains in employee performance that would be realized at different passing score levels.

Highlights of the research findings are summarized below.

Reading & Writing Test Scores Predict Academy and Job Performance

Overall Results

Consistent with previous validation research conducted by POST over the last decade, as well as published cumulative validity evidence for verbal ability tests in predicting performance in law enforcement occupations, Reading & Writing Test scores were found to be predictive of subsequent performance both in basic academy training and on the job.

More specifically, Reading & Writing total scores were predictive of *academy performance* as measured by instructor ratings of students' writing ability, overall success/failure in completing training, and Academy Proficiency Test scores. In addition, Read/Write scores were found to predict *job performance* as measured by FTO ratings of officers' writing ability demonstrated throughout field training. Total score correlations with overall success/failure in completing field training were *not* significant (p>.05) in the present study, although such performance was predicted in a previous POST study in which a larger and broader sample of officers was obtained (Weiner & Berner, 1987).

Each of the five subtests (Clarity, Spelling, Vocabulary, multiple-choice Reading Comprehension and Cloze) was found to predict rated performance in both basic training and field training, and all except one (Clarity) were predictive of overall academy success/failure.

Essay Test scores were predictive of *academy performance*, including instructor ratings of writing ability, overall success/failure and Proficiency Test scores. However, only certain elements of FTO ratings of officers' writing ability demonstrated in *field training* were predicted (Organization & Narrative and Mechanics), and overall field training success/failure was *not* predicted.

Within-Group Validity

Read/Write total scores were found to predict academy performance ratings and Proficiency Test scores within all racial/ethnic and gender groups studied; i.e., Asians, Blacks, Hispanics, Whites, males, and females; and for Proficiency Test scores only, American Indians and Filipinos. Within-group predictions of both overall academy success/failure and FTO ratings were not significant in many instances. However, the numbers of subjects within many of the subgroups were small, and thus, statistical power to detect significant correlations was often low.

Essay Test scores were found to predict *academy performance ratings* and *overall academy success/failure* within racial/ethnic and gender groups with few exceptions. Academy instructor ratings of writing ability were predicted for all groups studied except one (a positive, but nonsignificant correlation was obtained for Asians); and overall academy success/failure was predicted for all groups except two (a positive, but nonsignificant correlation was obtained for Blacks; and Asians were not studied due to the small number of subjects in this group). Proficiency Test scores were predicted within-sex but *not* within-race/ethnicity. Performance in *field training* was *not* predicted within the various groups in most cases.

Differential Prediction

In a number of instances, significantly different prediction equations were obtained for racial/ethnic minorities versus Whites, as well as for females versus males when measures of their academy performance and field training performance were regressed onto Reading & Writing Test scores and Essay Test scores. However, the net results of the racial/ethnic minority-majority group differences were found to be neutral or to actually favor the minority group; i.e., on average, their performance was either the same as would be predicted by their test scores, or was significantly *over*predicted by their test scores. In this sense, test scores were *not* found to be unfair to the racial/ethnic minority groups studied. The results for females were not consistent; in some instances their performance was *over*predicted and in others it was *under*predicted. However, females generally perform well on the Reading & Writing Tests, thus obviating concerns regarding differential impact.

Alternative Test Batteries Offer Little or No Gain in Prediction

Several alternative test batteries were constructed consisting of different configurations of the five Reading & Writing subtests and the Essay Test. The alternative batteries were assembled in consideration of: (a) individual test validities, (b) the relative difficulty of the tests for racial/ethnic and gender groups, (c) the joint relationships among the tests as predictors of the various criteria, and (d) representation of reading and writing abilities.

Overall, little or no significant gain in validity was realized by the alternative test

batteries relative to the current battery in predicting performance in either basic training or on the job. Moreover, no significant reductions in difficulty for racial/ethnic minorities were detected for any of the alternative batteries relative to the current battery. It is noteworthy that comparable prediction of academy and field training performance was achieved in some instances with fewer than five tests.

Moderator Variables Found to Have No Effect Upon Validity

Several variables were examined with respect to their effects upon the criterion measures of academy and field training performance and, in turn, the correlations obtained between test scores and these criteria. These variables included: (a) *time* between testing and criterion data collection; (b) potential sources of *rating inaccuracy or contamination*, such as the evaluator's degree of familiarity with the ratee's job performance, the officer's physical fitness/appearance, and personality traits exhibited by the officer; and (c) *rater characteristics*, including race/ethnicity and gender. Overall, the results indicated that none of the variables studied were found to moderate the obtained validity results. Moreover, there was no significant trend for FTOs in racial/ethnic and gender work force majority groups to give lower ratings to minority group members.

Reading & Writing Test Scores Offer Utility

Reading & Writing Test scores in the lower to upper middle range (35, 40, 45, 50 & 55) were examined with respect to their utility in making pass/fail employee selection decisions. Utility was expressed as the relative percentage gain in (a) *adequate* and (b) *above average* performing employees that would be realized if a given Read/Write score was used to make a pass/fail decision. Several expectancy tables were constructed demonstrating the relationships between Read/Write Test cut scores and different measures of performance. These results represent gains *relative to existing personnel selection practices*.

With regard to prediction of *academy instructor ratings*, Read/Write Test scores at each pass/fail level were found to yield significant gains in student performance, ranging from approximately 3% to 29% for *above average* performance, and 1% to 5% for *adequate* performance. Read/Write Test scores were also found to offer utility in predicting *Academy Proficiency Test scores*, with gains ranging from approximately 3% to 37% for *above average* performance, and 2% to 17% for *adequate* performance. Gains in officers' rated performance in *field training* were limited to the 45-55 score range, ranging from approximately 5% to 25% for *above average* performance, and 2% to 3% for *adequate* performance. The base rate for *overall academy success* was very high, thus there was little room for improvement in the prediction of this index of performance.

Within-group analyses indicated that Read/Write scores offer significant gains in both *adequate* and *above average* performance on the Academy Proficiency Test for all

racial/ethnic and gender groups studied. In addition, significant improvements in the selection of *above average rated* students in *basic training* were realized for Hispanics, Whites, males and females; gains for Blacks were also positive but not statistically significant, likely due to the small sample size.

Conclusions

The importance of verbal comprehension and expression (reading and writing) abilities for law enforcement work is well documented, both in terms of job analytic and empirical test validation research. From this standpoint, there is a clear rationale for assessing entry-level law enforcement candidates' reading and writing abilities.

The results of the present research in concert with the results of other research studies summarized above indicate that the POST Reading & Writing Test Battery provides a reliable and valid measure of an examinee's aptitude to perform writing-related activities in basic training, in subsequent field training, and even later as a tenured patrol officer. The research results also suggest that Read/Write Test scores are predictive of students' acquired knowledge of the basic training curriculum, as measured by achievement test scores. In addition, Read/Write scores were found to provide a measure of the likelihood of successful completion of basic training.

The results pertaining to alternative test batteries suggest that while little or no significant gain in prediction resulted from several alternative test batteries studied, it may be possible to construct a shorter test battery that predicts academy and job performance comparably to the current battery. Analyses of the relative difficulty of alternate test configurations suggests that further study may be warranted to explore the apparent underlying differences in the difficulty of the tests for various subgroups.

Finally, consideration should be given to the *criterion problem*. That is, despite the best efforts of researchers, it is difficult to obtain highly reliable and valid criterion measures of training performance and job performance. This is a chronic problem in personnel selection research and is widely acknowledged in the literature (e.g., Hirsh, et al., 1986). Implications of this problem for the results of the present study include potential underestimation of validity and utility, as well as inaccurate assessments of group differences. Thus, future directions for subsequent research would do well to include attempts to develop better criterion measures.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). <u>Standards for educational and</u> <u>psychological testing</u>. Washington, D.C.: American Psychological Association.
- Bernardin, H. J. (1988). Police officer (Chapter 10.8), in S. Gael (Ed.), <u>The job analysis</u> <u>handbook</u>. New York: Wiley & Sons.
- Berner, J. G., & Kohls, J. W. (1982). <u>Patrol officer physical performance testing manual</u>. Sacramento: Commission on POST.
- Berner, J. G., et al. (1985). <u>Analysis of the physical demands of California patrol officers</u>. Unpublished technical report. Sacramento: Commission on POST.
- Cascio, W. F. (1982). <u>Costing human resources: The financial impact of behavior in</u> <u>organizations</u>. Boston: Kent Publishing.
- Cohen, J., & Cohen, P. (1975). <u>Applied multiple regression/correlation analysis for the behavioral sciences</u>. New York: Lawrence-Erlbaum.
- Cohen, J. (1988). <u>Statistical power for the behavioral sciences</u> (p. 85). New Jersey: Lawrence Erlbaum.

Guilford, J. (1954). Psychometric methods (2nd ed.) (p. 393). McGraw-Hill.

- Guilford, J., & Fruchter, B. (1973). <u>Fundamental statistics in psychology and education</u> (5th ed.) (pp. 315-316 and 439-442). McGraw-Hill.
- Hirsh, H., Northrop, L., & Schmidt, F. (1986). Validity generalization results for law enforcement occupations. <u>Personnel Psychology</u>, 39.
- Honey, R. A., & Kohls, J. W. (1981). <u>Entry-level reading and writing tests for California</u> <u>law enforcement: Final report</u> (Technical Report No. 4). Sacramento: Commission on POST.
- Honey, R. A. (October, 1983). <u>User's manual for the POST Entry-Level Law Enforcement</u> <u>Test Battery</u>. Sacramento: Commission on POST.
- Kohls, J., Berner, J. & Luke, L. (1979). <u>California entry-level law enforcement officer job</u> <u>analysis</u> (Technical Report No. 1). Sacramento: Commission on POST.

Miner, M. G., & Miner, J. B. (1979). <u>Analysis of Uniform Guidelines on Employee Selection</u> <u>Procedures</u>. Washington, D.C.: The Bureau of National Affairs, Inc.

U.S. Department of Labor (1977). Dictionary of Occupational Titles (4th ed.).

- Weiner, J. A., & Berner, J. G. (1987). <u>POST Basic Course Proficiency Test study:</u> <u>Relationships with patrol officer job performance</u>. Sacramento: Commission on POST.
- Weiner, J. A. (1988). <u>Criterion-related validation study of the POST Work Sample Test</u> <u>Battery</u>. Sacramento: Cooperative Personnel Services.
- Weiner, J. A. (1993). <u>Physical abilities test follow-up validation study</u>. Sacramento: Commission on POST.

APPENDICES

APPENDIX A

SCORING PROCEDURE FOR THE

READING & WRITING TEST BATTERY



Scoring Procedure for the Reading & Writing Test Battery

1. Compute raw percent correct scores on each subtest; i.e., for each subtest, divide the number of correct responses by the number of items:

Clarity (Cl) Spelling (S) Vocabulary (V) Reading Comprehension, multiple-choice (RC) Cloze test (Cz)

2. Compute mean percent composite scores for Writing (W) and Reading (R):

$$W = (Cl + S + V) / 3$$

$$\mathbf{R} = (\mathbf{R}\mathbf{C} + \mathbf{C}\mathbf{z}) / 2$$

Compute Writing T score (Wt) and Reading T score (Rt) by calibrating each mean percent composite score to the means and standard deviations (SDs) obtained for a benchmark validation study sample (POST Entry-Level Law Enforcement Test Battery User's Manual, 1983).

Wt = ((W - 76.6) / 10.1) * 10 + 50

Rt = ((R - 67.2) / 11.7) * 10 + 50

4. Compute Total T score (T) by summing the Writing and Reading T scores, then calibrating the sum to the 1983 validation sample mean and SD, and then rescaling to a T scale:

T = ((Wt + Rt) - 99.9) / 17.8) * 10 + 50

The range of possible scores on the Reading & Writing Test Battery (T) is -25.1 to 78.6.

.

3.

APPENDIX B

ESSAY TEST INSTRUCTIONS AND SCORING PROCEDURE

a ;

ESSAY INSTRUCTIONS

You will have forty minutes to write a legible and complete essay on the following essay topic.

ESSAY TOPIC

Describe an event that made a change in your life. Explain why that event had importance for you.

SUGGESTIONS FOR TAKING THE TEST

Consider the topic carefully. Organize your response before you begin writing. Fit your response into the time allotted.

DEPARTMENT USE ONLY

Reader		Score	
Reader		Score	
Reader	·····	Score	
	FIN	AL SCORE	

ESSAY NO. <u>E-1</u>

POST SCORING GUIDE

Candidates should be rewarded for what they do well. They are asked, first, to narrate or describe an event or situation from personal experience. In the last part of the prompt, they are directed to provide some sort of analysis of the experience. Although the assignment calls for a two-part response, one part may be implicit in the other.

RANGE OF SCORES

- 6 The "6" essay will be fluent, well developed, and well organized. It will show clear command of language and will be relatively free of errors in sentence structure, grammar, and mechanics.
- 5 The "5" paper, despite occasional faults, will be generally well written and well organized. It will be less fluent and less detailed than the "6" paper, but will demonstrate greater facility than the "4" paper.
- 4 The "4" paper will demonstrate basic writing competence, though it may have some problems in sentence structure, diction, or mechanics or have limited development.
- 3 The "3" paper may not provide adequate development, may lack detail and specificity, or may be poorly organized. It usually has problems in diction, grammar, and mechanics.
- 2 The "2" paper may lack coherence or adequate development. Generally, it will be marred by multiple errors in sentence structure, grammar, and mechanics. It suggests incompetence.
- 1 The "1" paper will show clear incompetence.

Non-response papers and off-topic papers should be given to the chief reader.

5 4 A.

and the second second

APPENDIX C

BASIC ACADEMY REPORT WRITING INSTRUCTOR RATING BOOKLET END OF WRITING INSTRUCTION EVALUATION

化合金化合金属 医骨髓炎 化过度器 医结节 化苯乙酰氨基化 法管理法 法无法定法

Commission on Peace Officers Standards & Training Reading and Writing Test Study

REPORT WRITING INSTRUCTOR RATING BOOKLET END OF WRITING INSTRUCTION EVALUATION

Date:	//		
Academy:		 	_ _
Evaluator:		 	
Position at Academy:			· .

Using the rating scales contained in this booklet, you are to evaluate the performance of the cadets listed on the following pages. The ratings you provide will be used in a POST study to evaluate alternative entry-level reading and writing testing requirements. Your ratings will be used only for the purposes of this study, will have no bearing on the training or employment status of the cadets, and will be kept <u>confidential</u>. Please be completely candid and objective in making your ratings.

INSTRUCTIONS

You will be evaluating each cadet on four separate abilities that are necessary to write good reports. While these abilities are all required for good report writing, they are unique and different. Therefore, when making your evaluations, be sure that you pay particular attention to the specific ability under consideration.

In addition, for each writing ability there is a <u>different</u> 5-point rating scale. Thus, be sure to review carefully the definition of each scale point on each rating scale.

Frame of Reference. When rating each cadet's report writing abilities, be sure to evaluate the cadet's <u>current</u> competency to write police reports.

Avoid Common Rating Errors. Avoid common rating errors by following the below guidelines:

- 1. <u>Carefully consider each specific ability to be evaluated.</u> A common rating error, "Halo," occurs when the evaluator gives an individual the same or very similar ratings on a range of performance factors based on some global impression of the individual's performance. Avoid this error by carefully considering each ability separately when making your ratings.
- 2. <u>Use of the full range of the rating scale</u>. Another type of rating problem occurs when the rater uses only one or two points on the rating scale (i.e., rates everyone the same). Avoid this error by carefully considering each scale point on the various rating scales when making your ratings.
- 3. <u>Use the rating scales as defined</u>. A third common rating error occurs when the rater uses his or her own definition of the ability being evaluated, resulting in inaccurate ratings. Avoid this error by reviewing carefully the definition of each ability to be evaluated, as well as the descriptions of the scale points on each of the rating scales.

The four writing abilities that you will be evaluating are defined as follows:

Ability 1 - ORGANIZATION AND NARRATIVE: The ability to compose clear and organized narratives in reports.

Ability 2 - WRITING MECHANICS: The ability to write reports that are free of errors in fundamental writing mechanics (i.e., reports that are characterized by good grammar, punctuation, spelling, and word choice).

Ability 3 - INFORMATION AND ELEMENTS: The ability to include all necessary information and elements in reports.

Ability 4 - TIMELINESS: The ability to write acceptable reports in a timely manner.

Evaluate all cadets on a single ability before proceeding to the next ability.

Remember to evaluate each cadet's current report writing abilities.

POST READING & WRITING CRITERIA RESPONSE FORM

Using the rating scale below, decide which scale value best describes the reports written by each cadet you are evaluating. Record your evaluations in the spaces provided to the right.

Ability 1--ORGANIZATION AND NARRATIVE: The ability to write clear and organized narratives in reports.

ORGANIZATION AND NARRATIVE

- 5. <u>Excellent:</u> Reports are fluent, well developed, and well organized. They show clear command of language and are clearly and logically presented. No, or very little, editing is required to improve the narrative.
- 4. <u>Good:</u> Despite occasional faults, reports are generally well written and well organized. They are less fluent and less detailed than an excellent report, but demonstrate greater facility than an adequate report. Reports at this level require little, if any, editing to improve the clarity of the narrative.
- 3. <u>Adequate:</u> Reports demonstrate basic writing competence though they may contain problems in sentence structure or diction, or have limited development. Occasionally, some revision is required to ensure proper interpretation. This represents the minimum acceptable level of performance.
- 2. <u>Poor:</u> Reports often require further development, lack detail and specificity, or are poorly organized. Reports at this level typically require extensive revision and rewriting.
- 1. <u>Very Poor:</u> Reports lack coherence and/or adequate development. Reports at this level are not suitable for revision.

	Name	· .		Ability No. 1: Organization &
Last	First	M.I.	SSN	Narrative
	······································			
	<u>,</u>			
анан алан алан алан алан алан алан алан			and a second	
				-1
	<u>, 1999, 1999, 1999, 1999, 1999</u> , 1999, 1990, 1999, 19	<u></u>		
		· · · ·		

	Name	 Distanti di Contra di C			Ability No. 1: Organization &
Last	First	M.I.		SSN	Narrative
					ta ta si a dana da sa
	۵۰۰۰ <u>میلاد میلاد در این این اور اور اور اور اور اور اور اور اور اور</u>				
		<mark>nong an Padak Kanadara (Padar</mark> Kanadar (Padar) Manadar (Padar)			- <u>La construction construction por con</u>
	مى يې د مېزى د يې يې د مېزى يې د مېزى يې د مېزى يې د يې يې د مېزى يې د	<u></u>		an a	من المانية موجود ويوموسون ک
		<u>, and an and a state of the st</u>		<u>an an a</u>	
	فسر ومناقبة فالمعر والمعارف والمراجع والمراجع والمراجع	<u>a ta ang ang ang ang ang ang ang ang ang an</u>		<u>er 1990 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 19</u>	۰۰
		Lianza, Muumariin debuina terret waterinia			a na sa
				an and a support of the second se	in the state of the state
	<u>محمد میں ایک ایک میں پر یک این پر محمد ہوتا ہے۔</u> یہ ایک رواحا م <u>ے خاص میں پر ایک ایک میں ایک میں میں ایک میں ایک</u>	<u>an an an an Anna an Anna an Anna an Anna</u>			
المعني مانت شديدة المعنية عليه عنها السالي عن في مانية عليهم ومراشل والعمل والعمل. الم	<u> </u>			ىشىنىڭ سەتلەر ب ۆكەتلەردۇمىيە ئىلىپۇر يەرە يەيرىپىرى مەتتەپىرى بىرىيىتى يەرىپىرىيىي	and the second secon
<u>ىسى ئىلىدىكە بىرى بىرىنى ئەلىكە بىلى بىلى بىلى بىلى بىلى بىلى بىلى بىل</u>	<u>en an ana ing ing ang ang ang ang ang ang ang ang ang a</u>				<u>and the stand of the standard</u>
لىمارى بىلار بىرىغان بىلەر بىرىغان بىلەر بىرىن بىرىغان تىرىپ <u>مەرىپ بىرىمى بىرىپ بىرىپ بىرىپ بىرىپ بىرىپ بىرىپ</u>	an a	<u>n ang sang sang sang sang sang sang sang</u>		an a	and the second
					ويتفسقيسان بتر بريسيتن
				an a	يونون بويون رويون روي
			•		
مىمىيە ئەيەر مەللىلىغەنچىمە مەھىيەت بىرىيەت بەر يېلىغىنىيە يەر بەر ئېرىغانىيە بۇرىيەت ئېرىمەت ئېرىيە بىرىيە.	an a	<u>, - 12 - 12 - 20 - 20 - 20 - 20 - 20 - 20</u>			
and a second provide a second seco)	<u>an na shina na shina</u>	ۇيىلارلەركىيەتىرىنىڭ ئەلەركىيەتىيە بىلەرلەركىيەت ئەلەرلەركىيەت بىلەرلەركىيەت بىلەرلەركىيەت بىلەرلەركىيەت بىلەرلەركىيەت بىلەرلەركىيەت بىلەرلەركىيەت بىلەركىيەت بىل
	in a standard för til og som för det av som förstade standard för som	and a second	·	and a second	And the state of the
•	وزرت مسيع مسينا بالمجبر وارتباط وبالي مجتمرا وبرو بيور محيات			د	and the second
		ا مەلەر بىرا ئەتلەر قەرىپ بىرىسىدۇر يېرىسىدۇرىغ			ىرىنىيە ئېرىنىيە بېيىرىيە يېسىر <u>ت</u> ى
		م الم الم الم الم الم الم الم الم الم ال			an a
				•	
۱۹۹۵ - ۲۰۰۰ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ ۱					
<u>ى</u>			•	and a second	
				and a second	
مىلىمەت كەركىكى بىلىرىكى <u>بىلىرىكى بىلىرىكى بىلىرىكى بىلىرىكى بىلىرىكى بىلىرىكى بىلىرىكى بىلىرىكى بىلىرىكى بىل</u>	n a star and a star and a star and a star	<u></u>	•	۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹ - ۲۰۰۹	<u>an an a</u>
					and the state of the
		an an thair an		and the second	

Using the rating scale below, decide which scale value best describes the reports written by each cadet you are evaluating. Record your evaluations in the spaces provided to the right.

<u>ABILITY 2--WRITING MECHANICS: The ability to write reports that are free of errors in fundamental writing mechanics (i.e., reports that are characterized by good grammar, punctuation, spelling and word choice).</u>

WRITING MECHANICS

- 5. <u>Excellent:</u> Reports show a clear command of the language and generally contain very few, if any, errors in grammar, punctuation, spelling or word choice. Reports require no, or very little, editing to correct technical writing faults.
- 4. <u>Good:</u> Despite occasional technical writing faults, reports are generally well written and require limited editing.
- 3. <u>Adequate:</u> Reports demonstrate basic competence but usually contain some errors in grammar, punctuation, spelling or word choice. Reports at this level sometimes require revision to ensure proper interpretation. This represents the minimum acceptable level of performance.
- 2. <u>Poor:</u> Reports are marred by frequent errors in sentence structure, punctuation, spelling, or word choice. Problems with mechanics make editing for correctness extremely difficult.
- 1. <u>Very Poor:</u> Reports contain too many technical errors to correct. Reports at this level are not suitable for revision.

	Name		andra Alexandra Alexandra	Ability No. 2:		
Last	First	M.I.	SSN	Writing Mechanics		
and and a second se Second second second Second second						
						
**************************************			n - Antoine an Anna an Anna Anna Anna Anna Anna A	•		

an a	Name			SSN 1997	Ability No. 2: Writing Mechanics
Last	1150				
<u></u>					<u>eren sedera en esta</u>
				<u>an ang kang sana sana sana kang kang kang sana sana sana sana sana sana sana s</u>	<u>, a tradición de la compañsion</u>
				<u>an mananan seban yang persang ang bin pinaman tahun na bata</u> P	and the second second second
		<u>a de la composition de</u>		<u>al and a sub-sub-sub-sub-sub-sub-sub-sub-sub-sub-</u>	<u>man jarjabat oʻratlar moʻr</u>
		an a			and the product of the second s
		anna an an ann an ann an ann an an ann an a		a a a a a a a a a a a a a a a a a a a	and managements of contracts
		an a		a an	un an
					and a second
	a the second				
۵٬۰۰۰ میلاد میکند. ۱۹۹۹ - ۲۰۰۰ میلاد میکند با ۲۰۰۰ میلاد میکند میکند. ۱۹۹۹ - ۲۰۰۰ میلاد میکند میکند.	in and a second seco	and the set of the set			
	ann an tha tha an t	<u>an an a</u>		<u>an an a</u>	A before the cost of cost of the cost of
		<u>, a construir an ann an ann an an an an an an an an a</u>	· •	an a	and the second
		an a			ensisten werdensister?
	and a start of the s	and the second state of the se		and an an other manifest set in a second second set of the second second second second second second second se	with the refrequences
		an an an an ann an an an an an an an an		a na sana na sana na	
	ىرىمىيە تېرىكىيىرىكىيىتىنى دەر مىلىتى <u>مەرىمىيە مىلىتى مىلىتى مىلىتى</u>			an a	<u>an ar star e c'harren ar t'harren der t</u>
		-		and the second	and a strange for the state of the
					<u>, son sing site same same same same</u> sa
	<u></u>				
		<u>an an a</u>			
		<u>an an a</u>	. •	<u>A BARAN MANAN MANAN MANAN MANAN MANAN MANAN MANAN MANAN MANAN</u>	<u>an an a</u>
	<u></u>			<u> </u>	<u>. The specific states and the second states of specific states and the second states are specific states and the specific states are specific states and the specific states are specific states and the specific states are spec</u>
	ананан каландаган каландаган каландаган каландаган каландаган каландаган каландаган каландаган каландаган калан Тарак	<u>an an a</u>			in formation and an other and a second se
	,,,,	and the second	a na ang ang ang ang ang ang ang ang ang		and the second
	العنيان مرين المتراد مريني مسرور	<u></u>	• • •	an a	<u>(in principal constructions and a state state</u> of

, ,

POST READING & WRITING CRITERIA RESPONSE FORM

Using the rating scale below, decide which scale value best describes the reports written by each cadet you are evaluating. Record your evaluations in the spaces provided to the right.

ABILITY 3--INFORMATION AND ELEMENTS: The ability to include all necessary information and elements in reports.

INFORMATION AND ELEMENTS

- 5. <u>Excellent:</u> Reports contain all essential and relevant information as well as all elements. Information and elements are clearly and logically presented. No, or very little, editing is required.
- 4. <u>Good:</u> All essential information and elements are present, but reports may contain minor omissions of relevant information. Information and elements are not as clearly and logically presented as in excellent reports. Little editing is required of reports at this level.
- 3. <u>Adequate:</u> Essential information and elements are present, but there may be omissions of relevant information and the elements may not be clearly presented. Parts of the reports may have to be rewritten to ensure proper interpretation. Information and elements are presented just well enough to satisfy minimum requirements.
- 2. <u>Poor:</u> Essential information and elements are omitted. Reports at this level typically require extensive revision and rewriting.
- 1. <u>Very Poor:</u> Much necessary information and many elements are omitted. Reports at this level are not suitable for revision.

	Name		•		A	bility No. 3	
Last	First	M.I.		SSN	Elements		
					-		
		en e					

Last	First	M.I.	n de la constante de la constan La constante de la constante de	Addity N Info. & Eleme
			and a second	
		an a		
<u></u>				- الفاقر بينيين
والمتلا كالفلا والمتحاد والمحادث والمحادث				
				and the second
		<u>a ang kanang na sana ana s</u> a ang kanang na sana sa	alet and a second a	
	<u> 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997</u>			an a
	<u>, en </u>			
<u>an an a</u>				
				i sente a comunicación a comunicación de la comunicación d
				1
	en de la constante de la const La constante de la constante de	<u>na de antes series de la constante</u>		
				andra an Andra andra andr
<u>a: 1:2, -1:0 - 1:1:0 - 1:0 1:0 1:</u>	an da an		<u>anti-site and a second s</u>	
		a tanàna mandritra dia kaominina dia kaominina dia kaominina dia kaominina dia kaominina dia kaominina dia kaom Ikang kaominina dia kaominin	na sense a sen A sense a sense A sense a sense	
			kan ang kanalan ang kanala Ang kanalang	· · · · · · · · · · · · · · · · · · ·
	1997년 - 1997년 - 1997년 - 1997년 1998년 - 1997년 - 1997년 1997년 - 1997년 -			
<u></u>				
<u>alaan maring an maring a sana</u>			an a	
				<u>in terintenden die die d</u>
			Start Street	¢. Sur

, b.,

<u>ر</u> .

POST READING & WRITING CRITERIA RESPONSE FORM

Using the rating scale below, decide which scale value best describes the reports written by each cadet you are evaluating. Record your evaluations in the spaces provided to the right.

ABILITY 4--TIMELINESS: The ability to write acceptable reports in a timely manner.

TIMELINESS

- 5. Excellent: Always produces accurate, well written reports in less time than is typical.
- 4. Good: Often produces accurate, well written reports is less time than is typical.
- 3. <u>Adequate:</u> Usually produces accurate, well written reports within a reasonable period of time.
- 2. <u>Poor:</u> Often requires an excessive amount of time to produce an accurate, well written report.
- 1. <u>Very Poor:</u> Always requires an excessive amount of time to produce an accurate, well written report.

	Name		Ability No. 4:
Last	First M.I.	SSN	Timeliness
			
7			

	Name	a Maria Maria Maria Maria Maria Maria Maria Maria Maria Maria		Ability No. 4
Last	First	M.I.	SSN	Timelines
			and a start of the second s Second second	and a second
				ka di kacamatan kacam Manang kacamatan kacam
			en e	
				a di seri di s
	<u>enterne en enterne in en en</u>			
				-
				• • • • • • • • • • • • • • • • • • •
	<u></u>			
			an a	
	1			
<u>.</u>				
		<u></u>		en la construcción de la
<u></u>				
			and a start and A start and a st A start and a st	
M				
		A	на селото с селото с на селото с	
				_
• •****				
. ?				

APPENDIX D

BASIC ACADEMY SUCCESS/FAILURE

DATA COLLECTION FORM AND CODING INSTRUCTIONS

an an an the second and a second second

POST READING AND WRITING TEST AND PHYSICAL TEST RESEARCH DATA FORM

tata a annuale tam ter each annter di

ACADEMY	~		CLAIR HUGHEN						
START DATE OF ACADEMY		END DATE OF ACAD	END DATE OF ACADEMY					•	
Complete this section at the beginning of the academy class.						Comple	te this a of the d	ection at f	the completion
Name			Status at beginning of training (indicate one of the following) 1. Employed as full-time officer 2. Reserve officer 3. No effiliation		Training outcome: See reverse side of this form for codes				
Lest,	First	Middle I.	Social Security Number	•	If employed or a reserve name of employing agency				
•							1.1		
									2000 - 100 100 - 100 100 - 100
	·								
						 			
							ļ		

····							ļ		
							 		
						ļ	 		
	·····			ļ		 			
				L		1			• • •

CODES FOR ACADEMY TRAINING OUTCOME

COMPLETED TRAINING:

- C1 = Graduated academy in normal time.
- C2 = Graduated academy but required extra time (remediation).

FAILED TO COMPLETE TRAINING:

RESIGNED (VOLUNTARY)

R1 = Overall academy performance was satisfactory.

Unsatisfactory academy performance due to ...

- R2 = inadequate report writing skills.
- R3 = generally inadequate analytical skills.
- R4 = inadequate weaponless defense skills.
- R5 = inadequate baton skills.
- R6 = failure to complete POST physical conditioning program or Work Sample Test Battery.
- R7 = inadequate knowledge, skills and abilities other than report writing, analytical and physical (e.g., knowledge of laws, procedures, tactics; learning ability; oral communication skills; driving skills; etc.).
 - Cathor as a stitute matintian work babits)

R8 = Other reasons (e.g., attitude, motivation, work habits)

TERMINATED (INVOLUNTARY)

T1 = Overall academy performance was <u>satisfactory</u>.

- Unsatisfactory academy performance due to ...
- T2 = inadequate report writing skills.
- T3 = generally inadequate analytical skills.
- T4 = inadequate weaponless defense skills.
- T5 = inadequate baton skills.
- T6 = failure to complete **POST physical conditioning program or Work Sample Test Battery**.
- T7 = inadequate knowledge, skills and abilities other than report writing,
 - analytical and physical (e.g., knowledge of laws, procedures, tactics;

learning ability; oral communication skills; driving skills; etc.).

T8 = Other reasons (e.g., attitude, motivation, work habits)

INJURY

I1 = Cadet withdrew because of an injury.

RECYCLED

Cadet was recycled to attend next academy due to...

- N1 = injury or illness.
- N2 = inadequate report writing skills.
- N3 = generally inadequate analytical skills.
- N4 = inadequate weaponless defense skills.
- N5 = inadequate baton skills.
- N6 = failure to complete POST physical conditioning program or Work Sample Test Battery.
- N7 = inadequate knowledge, skills and abilities other than report writing,

analytical and physical (e.g., knowledge of laws, procedures, factics;

learning ability; oral communication skills; driving skills; etc.).

N8 = Other reasons (e.g., family emergency)

**IMPORTANT NOTE: More than one reason for failure to complete academy training may be coded.

APPENDIX E

FTO RATING BOOKLET

FINAL EVALUATION

le i kalen gentale ave Novi kalende og her

Commission on Peace Officer Standards & Training Physical Test and Reading & Writing Test Study

FTO RATING BOOKLET

FINAL EVALUATION

Date:

Agency: Los Angeles PD

(19420)

FTO	/Eval	lator

Name: _____

SSN:

Trainee	<u>1</u>		
Name:		 	
SSN:			

This booklet contains 3 parts. Part 1 is a practice rating exercise designed to help standardize the ratings made by you and other FTO's. Part 2 calls for you to rate the importance of various aspects of patrol work. In part 3 you are to rate the performance of the above designated trainee throughout the field training period. The ratings you provide will be used in a POST study to follow-up on the job performance of basic academy graduates. Your ratings will be used only for purposes of this study, will have no bearing on the employment status of the individual being rated, and will be kept **confidential** -- they will be forwarded directly to POST.

Thank you for your assistance.

Part 1

RATING EXERCISE

EVALUATOR INFORMATION

The following information is requested in order to document the representativeness of the FTO's participating in this study.



2

RATING EXERCISE

The following exercise was designed to do two things. One is to provide practice using the job performance rating scales contained in this booklet. The other is to provide examples of "good" and "poor" job performance to help guide you in evaluating the trainee's physical job performance.

In this exercise, you are to suppose that you have witnessed "officer X" (<u>a hypothetical</u> <u>officer</u>) performing various job activities. You are then asked to review and rate 24 different examples of job behavior exhibited by this fictional officer. You will be given feedback to show you how your ratings compare to the average ratings made by a sample of sergeants and field training officers. As you complete the exercise, **please keep in mind that there are no right or wrong answers**.

Procedure

1. <u>Review and evaluate each example job behavior</u> on the next two pages. Use the 5-point rating scale at the top of each page to indicate how you would rate the performance of an officer (any officer) who did this on the job.

You should consider each example behavior at "face value." Do not assume complicated or special circumstances.

For example, consider the first example job behavior on the next page. If you feel that an officer who does this on the job is performing at a "Very Poor level, Far Below Job Demands," then you should rate that behavior as a "1" on the scale.

- 2. <u>Complete your ratings</u> by writing the corresponding number next to each example.
- 3. <u>Compare your ratings</u>. As you rate the example behaviors, you will be periodically directed to different pages of this booklet to compare your ratings to average ratings for these same examples made by other FTO's and sergeants. Hopefully, you will find your ratings to be similar to those made by others. (<u>Note</u>: the average ratings are shown to the nearest tenth for your information only. Your ratings are to be a whole number, **1 to 5 only**).

Please do not change your ratings after making comparisons.

Now begin the rating exercise on the next page.

RATING EXERCISE

清洁:"通过关于主要们"为 eX

Use the following rating scale to indicate how you would rate an officer's job performance in each of the examples below.

VER	Y POOR POOR ADEQUATE GOOD EXCELLENT 1 2 3 4 5
Far Jol	Below Just Meets Far Exceeds Demands Job Demands Job Demands
How wou	d you rate an officer who:
1.	Starts to run after suspect but is exhausted within seconds; unable to continue foot pursuit.
2.	Has minor problems jumping obstacles but is usually able to maintain foot pursuit of suspect.
3.	Sweeps suspect to ground during struggle using hands and feet to control suspect.
4.	Drags and carries two children from a burning house.
*** NOW	TURN TO PAGE 5 AND COMPARE YOUR RATINGS TO EXAMPLE SET #1.
5.	Fails to maintain physical control of suspect and is hit with handcuff.
6.	Uses belt to tie suspect's feet to the car door when suspect tries to kick officer, after being handcuffed.
7.	Pursues suspect two blocks, jumps over a low fence and catches suspect.
8.	Has sustained several on-the-job injuries due to improper restraining techniques.
9.	Crawls through window in overturned vehicle to rescue accident victim.
10.	Tries to use baton and is overpowered by suspect, who takes the baton away from the officer.
.11.	Chases suspect on foot, climbs three walls and captures him.
12.	Is unable to restrain violent 5150 subject with leg irons.
*** NOW	TURN TO PAGE 6 AND COMPARE YOUR RATINGS TO EXAMPLE SET #2.

(continued)

RATING EXERCISE

VERY POOR	POOR	ADEQUATE	GOOD	EXCELLENT
1	2	3	4	5
Far Below	******	Just Meets		Far Exceeds
Job Demands	S	Job Demands	· · · ·	Job Demands

How would you rate an officer who:

- 13. Has no difficulty in kicking open locked door while pursuing 211 suspect on foot.
- 14. Misuses feet in self-defense causing undue injury to suspect.
- 15. Has no difficulty applying arm lock to subdue resisting suspect.
- 16. Runs after suspect who fied from stopped vehicle; after approximately 100 yards, catches and arrests suspect.
- 17. Is unable to lift or drag an accident victim from vehicle.
- 18. Uses baton to disable suspect by hitting his knee cap, causing him to fall down.
- 19. Is unable to crawl in confined areas quickly; is slow and unsure.
- 20. Jogs after suspect just fast enough to keep him in sight.
- 21. Loses balance easily and often falls while pursuing suspects on foot.
- 22. Chases suspect across a high and narrow catwalk, and catches the suspect.
- 23. Is unable to climb over wall, up to roof, allowing suspect to escape.
- 24. Properly handcuffs suspect who is kicking, biting and spitting.

*** NOW TURN TO PAGE 6 AND COMPARE YOUR RATINGS TO EXAMPLE SET #3

SET #1: (1) 1.0 (2) 2.6 (3) 3.3 (4) 4.6

AVOID COMMON RATING PROBLEMS

When making your performance evaluations, you can avoid some typical rating errors by following the guidelines below.

- Rate the officer's performance on each work component separately. Carefully consider each specific aspect of job performance to be evaluated. A common rating error, "Halo", occurs when the evaluator gives an officer the same rating in all areas of work because of a general impression of the officer's job performance.
- Use the full range of the rating scale. Another type of rating problem occurs when a rater adopts a rigid policy when making evaluations. For example, some raters may feel that no officers deserve to be rated very high (the rater is very strict), or that no officers should be rated very low (the rater is very lenient). Other raters may tend to "play it safe" by giving all subordinates an average rating.
- <u>Use the rating scales as they are defined</u>. Review carefully the specific definitions of each work component to be evaluated. A common rating problem occurs when raters simply read the titles and use their own definitions of the job components to be evaluated, resulting in inaccurate ratings.

This concludes the rating exercise. Please proceed to the next section of the booklet.

SET #2: (5) 1.3 (6) 3.9 (7) 3.2 (8) 1.3 (9) 3.6 (10) 1.3 (11) 3.7 (12) 1.5

SET #3: (13) 3.1 (14) 1.0 (15) 3.0 (16) 3.1 (17) 1.3 (18) 3.0 (19) 1.6 (20) 2.9 (21) 1.4 (22) 3.7 (23) 1.0 (24) 3.2 Part 2

IMPORTANCE RATINGS

PATROL OFFICER JOB ELEMENTS

Using the rating scale at the top of each page in this section, rate the importance of each listed item for successful performance as a patrol officer in your department. Mark your ratings in the space provided to the left of each item.

PHYSICAL JOB TASKS

IMPORTANCE SCALE

How important is competent performance of this task to the overall successful job performance of patrol officers in your department?

- 5 - Critically important
- 4 - Very important
- 3 - Important
- 2 - Of some importance
- 1 - Of little importance
- 0 - Not part of the job/ Unimportant for officers in my department
- 1. <u>Running</u> (e.g., pursuing suspects on foot; providing or obtaining emergency assistance).
- 2. <u>Handcuffing</u> suspects or prisoners (e.g., when apprehending and controlling subjects).
- 3. <u>Using restraining devices</u> other than handcuffs (e.g., leg irons, straps) to control subjects.
- 4. <u>Using baton</u> (or "Nun Chuku") to subdue attacking persons.
- 5. <u>Using locks, grips or holds</u> to subdue resisting persons (without using mechanical devices).
- 6. <u>Self-defense</u>, using hands or feet.
- 7. <u>Using body force</u> to gain entrance through barriers (e.g., locked doors) in routine and emergency situations (e.g., providing emergency assistance; investigating).
- 8. <u>Lifting/Carrying</u> disabled persons, equipment, heavy objects, etc., in routine and emergency situations where <u>speed</u> is often critical (e.g., providing emergency assistance; assisting the public).
 - 9. <u>Dragging/Pulling</u> disabled persons, equipment, heavy objects, etc., in routine and emergency situations where <u>speed</u> is often critical (e.g., providing emergency assistance; assisting the public).
 - 10. <u>Climbing</u> through openings (e.g., windows), over obstacles (e.g., walls), or up to elevated surfaces (e.g., roof) in routine and emergency situations where <u>speed</u> is often critical (e.g., pursuing suspects; providing emergency assistance; investigating).

8

IMPORTANCE SCALE

How important is competent performance of this task to the overall successful job performance of patrol officers in your department?

- 5 - Critically important
- 4 - Very important
- 3 - Important
- 2] Of some importance
- 1 - Of little importance
- 0 - Not part of the job/ Unimportant for officers in my department
- 11. <u>Crawling</u> in confined areas (e.g., attics) in routine and emergency situations where <u>speed</u> is often critical (e.g., providing emergency assistance; investigating).
- 12. <u>Jumping</u> over obstacles, down from elevated surfaces, or across openings in routine and emergency situations where <u>speed</u> is often critical (e.g., pursuing suspects; providing emergency assistance; investigating).
- 13. <u>Balancing</u> self on uneven or narrow surfaces (e.g., running up stairs; walking on building ledge; etc.) in routine and emergency situations (e.g., pursuing suspects; investigating; providing emergency assistance).

9

14. <u>Pushing heavy objects.</u>

(continued)

PHYSICAL ABILITIES

IMPORTANCE SCALE

1.

2.

How important is this ability to the overall successful job performance of patrol officers in your department?

- 5 - Critically important
- 4 Very important
- 3 - Important
- 2 Of some importance
- 1 Of little importance
- 0 - Not part of the job/ Unimportant for officers in my department
- <u>STRENGTH</u>: Exert physical force required to perform job activities (e.g., physically restraining others; lifting, pulling, pushing, or dragging hard-to-move objects; etc.).
- <u>ENDURANCE</u>: Maintain strenuous physical activity over prolonged periods of time (e.g., running long distance to pursue suspects; physically controlling resisting subjects; etc.).
- 3. <u>COORDINATION/AGILITY</u>: Move quickly and under control with rapid changes of direction, integrating the actions of arms and legs as required to perform job activities (e.g., running and jumping over obstacles, etc.).
- 4. <u>FLEXIBILITY</u>: Bend, extend and twist body segments as required to perform job activities (e.g., searching suspects, vehicles, buildings, etc.).
- 5. <u>OVERALL PHYSICAL ABILITY</u>: Perform the full range of physical job activities (e.g., pursuing suspects on foot; apprehending and controlling resisting/attacking subjects; providing emergency assistance; etc.).

JOB ELEMENTS

IMPORTANCE SCALE

1.

4.

How important is this job element to the overall successful job performance of patrol officers in your department?

5 -	Critically important
4 -	Very important
3 -	Important
2 -	Of some importance
1-	Of little importance
0 -	Not part of the job/ Unimportant for officers in my department

- <u>JUDGMENT</u>: Apply knowledge and reasoning to make prompt and effective decisions quickly in both routine and non-routine (e.g., life and death) situations; evaluate alternative courses of action and select the most acceptable alternative; make sound decisions in a timely manner; size up a situation quickly and take appropriate action.
- 2. <u>OBSERVATION SKILLS</u>: Recognize conditions or circumstances that indicate something might be wrong, unusual or out of the ordinary.
- 3. <u>LEARNING</u>: Comprehend new information and apply that which has been learned on the job.
 - <u>ORAL COMMUNICATION</u>: Speak in a clear, understandable manner and comprehend various types of information (e.g., accounts of past events, directions, explanations, ideas, etc.); talk effectively with persons of divergent cultural and educational background; speak with good pronunciation; project voice clearly; effectively use police radio.
- 5. <u>WRITTEN COMMUNICATION</u>: Write clearly and concisely; use acceptable grammar, punctuation and spelling; write reports that are well organized, complete and accurate.
- 6. <u>INTERPERSONAL BEHAVIOR</u>: Be sensitive to the feelings of others; resolve problems in ways that do not arouse antagonism; interact and deal effectively with people from varying social and cultural backgrounds; courteous and respectful; calm emotional people and attempt to resolve conflicts through persuasion rather than force.

JOB ELEMENTS (Continued)

IMPORTANCE SCALE

7.

How important is this job element to the overall successful job performance of patrol officers in your department?

- 5 Critically important
 4 Very important
 3 Important
 2 Of some importance
 1 Of little importance
 0 Not part of the job/ Unimportant for officers in my department
- <u>TEAMWORK</u>: Establish and maintain effective working relationships with co-workers, supervisors and other law enforcement officials (e.g., sharing information and working cooperatively with others, complying with departmental rules and regulations, following orders, accepting advice and constructive criticism, etc.).
- <u>ASSERTIVENESS</u>: Assert self when necessary to exert control over others; confront people who are behaving in a suspicious manner.
- <u>EMOTIONAL SELF-CONTROL</u>: Maintain composure and perform effectively in stressful situations; refrain from over-reacting when subjected to physical or verbal abuse; exercise restraint and use the minimum amount of force necessary to handle a given situation.
- 10. <u>ADAPTABILITY</u>: Adapt to changes in working conditions (e.g., changes in patrol assignment, shift changes, different types of incidents that must be handled one right after the other, etc.).
 - 11. <u>INITIATIVE</u>: Proceed on assignments without waiting to be told what to do; make an effort to improve skills and keep informed of new developments in the field; exert the effort needed to make sure the job is done correctly; consistently productive.
 - 12. <u>DEPENDABILITY</u>: Be reliable, thorough, punctual, accurate; assume responsibility for share of the workload; work with minimal supervision.
 - 13. <u>APPEARANCE</u>: Present a neat, clean, well-groomed appearance.

JOB ELEMENTS (Continued)

IMPORTANCE SCALE

How important is this job element to the overall successful job performance of patrol officers in your department?

- 5 - Critically important
 4 - Very important
 3 Important
 2 Of some importance
 1 Of little importance
 0 - Not part of the job/ Unimportant for officers in my department
- 14. <u>PHYSICAL FITNESS</u>: Maintain physical condition and fitness (e.g., exercise regularly; stay within reasonable weight limits).
- 15. <u>OFFICER SAFETY</u>: Apply appropriate safety precautions in hazardous and potentially dangerous situations; maintain a safe position when dealing with suspects; maintain awareness of location of self and others.
- 16. <u>LEGAL KNOWLEDGE</u>: Demonstrate working knowledge of laws, codes, and legal procedures (e.g., accurately detect crimes and violations and apply all appropriate codes; comply with legal requirements when making arrests, conducting searches, and obtaining evidence; write reports that include all necessary legal elements).
- 17. <u>KNOWLEDGE OF DEPARTMENT POLICIES AND PROCEDURES</u>: Demonstrate working knowledge of department policies, regulations and procedures (e.g., able to verbalize and apply them appropriately).
- 18. <u>KNOWLEDGE OF PATROL PROCEDURES</u>: Demonstrate working knowledge of procedures and techniques for performing patrol activities (e.g., able to verbalize and apply appropriate methods for beat patrol, suspect approach, vehicle stop, searching, restraining, prisoner transportation, and handling different types of calls).
- 19. <u>KNOWLEDGE OF INVESTIGATIVE PROCEDURES</u>: Demonstrate working knowledge of procedures and techniques for gathering information (e.g., able to verbalize and apply appropriate methods for locating and identifying victims, witnesses, and suspects; interviewing; collecting and preserving evidence).

· • 出版的 医哈尔二氏的脊髓上成的脊髓上的

的现在分词是不能是有一个人的人的是一个人的人的人,也是这个人的人的人,也是是我们就是我的人的人。 这些我们不可以有一个人的人的人,这次们是是这个人的

This concludes the importance rating section. Now proceed to Part 3 -- Final evaluation of Trainee Performance.

4. 人名法法法法法法 计算机算法 网络索尔兰 医尿道尿道 Part 3

FINAL EVALUATION of Field Trainee Performance

Using the rating scales contained in this section, evaluate the job performance of the trainee designated on the cover of this booklet. Please be completely candid and objective in making your ratings. Your responses will be kept confidential.

man have been and the man have been a set of the set of the

RATING INSTRUCTIONS

The performance appraisal is divided into four sections:

<u>Section I:</u> covers performance of physical job tasks and demonstrated underlying physical abilities;

Section II: covers performance of other (non-physical) job elements;

Section III: covers the abilities necessary to write reports; and

Section IV: covers the officer's overall job performance.

Section 1: Physical Job Performance

Steps to Follow in Making Performance Ratings:

1. Review the definition of the first job task on the next page.

2. Consider the trainee's job performance throughout the field training program, focusing on the job task to be evaluated.

Recall instances when the trainee's performance was especially good and/or especially poor.

- 3. Select the value from the **"Task Performance Scale**" below that best describes the trainee's performance of the task.
- 4. Write your rating in the space provided next to the task.
- 5. Proceed to the next task and repeat steps 1-4 above until you have rated the trainee's performance for all tasks listed.

Remember to avoid making common rating errors such as "halo" and "leniency".

TASK PERFORMANCE SCALE

How effective is the trainee in performing this task?

- 5 | Excellent: always extremely effective in performing this task.
- 4 - Good: performs this task effectively with little or no difficulty.
- 3 <u>Adequate</u>: performs this task just well enough to meet minimum job requirements.
- 2 | Poor: often has difficulty performing this task acceptably.
- 1 - Very Poor: unable to perform this task acceptably.
- N = Not observed/Unable to Rate

Section I-A: JOB TASKS

How effective is the trainee in performing job activities that involve ...

- 1. **<u>Running</u>** (e.g., pursuing suspects on foot; providing or obtaining emergency assistance).
- 2. <u>Handcuffing</u> suspects or prisoners (e.g., when apprehending and controlling subjects).
- 3. <u>Using restraining devices</u> other than handcuffs (e.g., leg irons, straps) to control subjects.
- 4. <u>Using baton</u> (or "Nun Chuku") to subdue attacking persons.
- 5. <u>Using locks, grips or holds</u> to subdue resisting persons (without using mechanical devices).
- 6. <u>Self-defense</u>, using hands or feet.
- 7. <u>Using body force</u> to gain entrance through barriers (e.g., locked doors) in routine and emergency situations (e.g., providing emergency assistance; investigating).
- 8. <u>Lifting/Carrying</u> disabled persons, equipment, heavy objects, etc., in routine and emergency situations where <u>speed</u> is often critical (e.g., providing emergency assistance; assisting the public).
- 9. <u>**Dragging/Pulling**</u> disabled persons, equipment, heavy objects, etc., in routine and emergency situations where <u>speed</u> is often critical (e.g., providing emergency assistance; assisting the public).
- 10. <u>**Climbing**</u> through openings (e.g., windows), over obstacles (e.g., walls), or up to elevated surfaces (e.g., roof) in routine and emergency situations where <u>speed</u> is often critical (e.g., pursuing suspects; providing emergency assistance; investigating).
- 11. <u>**Crawling**</u> in confined areas (e.g., attics) in routine and emergency situations where <u>speed</u> is often critical (e.g., providing emergency assistance; investigating).
- 12. <u>Jumping</u> over obstacles, down from elevated surfaces, or across openings in routine and emergency situations where <u>speed</u> is often critical (e.g., pursuing suspects; providing emergency assistance; investigating).
- 13. <u>Balancing</u> self on uneven or narrow surfaces (e.g., running up stairs; walking on building ledge; etc.) in routine and emergency situations (e.g., pursuing suspects; investigating; providing emergency assistance).
- 14. **Pushing** heavy objects.

Section I-B: PHYSICAL ABILITIES

Use the rating scale below to evaluate the trainee's physical abilities -- abilities which underlie the performance of various physical job activities. Follow the same procedures as outlined for Section I-A.

PHYSICAL ABILITY RATING SCALE

What level of ability does the trainee demonstrate on the job?

VERY POOR	POOR	ADEQUATE	GOOD	EXCELLENT
1	2	3	4	5
Far Below Job Demands	Often Unable to Meet Job Demands	Just Meets Job Demands	Often Exceeds Job Demands	Far Exceeds Job Demands

N = Not Observed/Unable to rate

PHYSICAL ABILITIES

- 1. <u>STRENGTH</u>: exerts physical force required to perform job activities (e.g., physically restraining others; lifting, pulling, pushing, or dragging hard-to-move objects; etc.).
- 2. **ENDURANCE**: maintains strenuous physical activity over prolonged periods of time (e.g., running long distance to pursue suspects; physically controlling resisting subjects; etc.).
- 3. <u>COORDINATION/AGILITY</u>: moves quickly and under control with rapid changes of direction, integrating the actions of arms and legs as required to perform job activities (e.g., running and jumping over obstacles, etc.).
- 4. **FLEXIBILITY:** bends, extends and twists body segments as required to perform job activities (e.g., searching suspects, vehicles, buildings, etc.).
 - 5. <u>OVERALL PHYSICAL ABILITY</u>: performs the full range of physical job activities (e.g., pursuing suspects on foot; apprehending and controlling resisting/attacking subjects; providing emergency assistance; etc.).

18

Section II: ADDITIONAL JOB ELEMENTS

This section of the rating booklet contains 18 job elements (abilities, skills, knowledge and behavioral traits) covering additional aspects of patrol officer work. Use the 5-point rating scale shown below to indicate the trainee's performance level on each job element. Apply the same rating steps as outlined in Section I.

JOB ELEMENT RATING SCALE

What level of performance does the trainee demonstrate on this job element?

VERY POOR	POOR	ADEQUATE	GOOD	EXCELLENT
	2	3	4	5
Far Below Job Demands	Often Unable to Meet	Just Meets Job Demands	Often Exceeds	Far Exceeds Job Demands
n an	Job Deman	ds	Demands	

N = Not Observed/Unable to rate

JOB ELEMENTS:

- <u>JUDGMENT</u>: Applies knowledge and reasoning to make prompt and effective decisions quickly in both routine and non-routine (e.g., life and death) situations; evaluates alternative courses of action and selects the most acceptable alternative; makes sound decisions in a timely manner; sizes up a situation quickly and takes appropriate action.
- 2. <u>OBSERVATION SKILLS</u>: Recognizes conditions or circumstances that indicate something might be wrong, unusual or out of the ordinary.
- 3. <u>LEARNING</u>: Comprehends new information and applies that which has been learned on the job.
- 4. <u>ORAL COMMUNICATION</u>: Speaks in a clear, understandable manner and comprehends various types of information (e.g., accounts of past events, directions, explanations, ideas, etc.); talks effectively with persons of divergent cultural and educational background; speaks with good pronunciation; projects voice clearly; effectively uses police radio.

Section II (cont'd): ADDITIONAL JOB ELEMENTS

VERY	POOR	POOR	ADEQUATE	GOOD	EXCELLENT
		2	radio contra destruct 	1997 - 1997 -	5
Far Jot	Below Demands	Often Unable to Meet Job Demands	Just Meets Job Demands	Often Exceeds Job Demands	Far Exceeds Job Demands

What level of performance does the trainee demonstrate on this job element?

N = Not Observed/Unable to rate

- 5. **INTERPERSONAL BEHAVIOR**: Is sensitive to the feelings of others and resolves problems in ways that do not arouse antagonism; interacts and deals effectively with people from varying social and cultural backgrounds; is courteous and respectful; calms emotional people and attempts to resolve conflicts through persuasion rather than force.
- 6. **TEAMWORK**: Establishes and maintains effective working relationships with co-workers, supervisors and other law enforcement officials (e.g., sharing information and working cooperatively with others, complying with departmental rules and regulations, following orders, accepting advice and constructive criticism, etc.).
- ASSERTIVENESS: Asserts self when necessary to exert control over others; confronts people who are behaving in a suspicious manner.
- 8. <u>EMOTIONAL SELF-CONTROL</u>: Maintains composure and performs effectively in stressful situations; refrains from over-reacting when subjected to physical or verbal abuse; exercises restraint and uses the minimum amount of force necessary to handle a given situation.
- 9. <u>ADAPTABILITY</u>: Adapts to changes in working conditions (e.g., changes in patrol assignment, shift changes, different types of incidents that must be handled one right after the other, etc.).
- 10. **INITIATIVE**: Proceeds on assignments without waiting to be told what to do; makes an effort to improve skills and keeps informed of new developments in the field; exerts the effort needed to make sure the job is done correctly; is consistently productive.

20

Section II (cont'd): ADDITIONAL JOB ELEMENTS

What level of performance does the trainee demonstrate on this job element?

VERY POOR	POOR	ADEQUATE	GOOD	EXCELLENT
1	2 2 2	1997 - 19	4	5
Far Below Job Demands	Often Unable	Just Meets Job Demands	Often Exceeds	Far Exceeds Job Demands
	to Meet Job Demar	lds	Job Demands	

N = Not Observed/Unable to rate

- 11. <u>**DEPENDABILITY</u>**: Is reliable, thorough, punctual, accurate; assumes responsibility for share of the workload; works with minimal supervision.</u>
- 12. <u>APPEARANCE</u>: Presents a neat, clean, well-groomed appearance.
- 13. **PHYSICAL FITNESS**: Maintains physical condition and fitness (e.g., exercises regularly; stays within reasonable weight limits).
- 14. **OFFICER SAFETY**: Applies appropriate safety precautions in hazardous and potentially dangerous situations; maintains a safe position when dealing with suspects; maintains awareness of own location and location of other officers.
- 15. <u>LEGAL KNOWLEDGE</u>: Demonstrates working knowledge of laws, codes, and legal procedures (e.g., accurately detects crimes and violations and applies all appropriate codes; complies with legal requirements when making arrests, conducting searches, and obtaining evidence; writes reports that include all necessary legal elements).
- 16. <u>KNOWLEDGE OF DEPARTMENT POLICIES AND PROCEDURES</u>: Demonstrates working knowledge of department policies, regulations and procedures (e.g., is able to verbalize and apply them appropriately).
- 17. <u>KNOWLEDGE OF PATROL PROCEDURES</u>: Demonstrates working knowledge of procedures and techniques for performing patrol activities (e.g., is able to verbalize and apply appropriate methods for beat patrol, suspect approach, vehicle stop, searching, restraining, prisoner transportation, and handling different types of calls).
- <u>KNOWLEDGE OF INVESTIGATIVE PROCEDURES</u>: Demonstrates working knowledge of procedures and techniques for gathering information (e.g., is able to verbalize and apply appropriate methods for locating and identifying victims, witnesses, and suspects; interviewing; collecting and preserving evidence).

Section III: WRITING ABILITY

In this section you will be evaluating each trainee on four separate abilities that are necessary to write good reports. While these abilities are all required for good report writing, they are unique and different. Therefore, when making your evaluations, be sure that you pay particular attention to the specific ability under consideration.

In addition, for each writing ability there is a <u>different</u> 5-point rating scale. Thus, be sure to review carefully the definition of each scale point on each rating scale.

When rating each trainee's report writing abilities, be sure to evaluate the trainee's <u>current</u> competency to write police reports.

The four writing abilities that you will be evaluating are defined as follows:

ABILITY 1 - ORGANIZATION AND NARRATIVE: The ability to compose clear and organized narratives in reports.

ABILITY 2 - WRITING MECHANICS: The ability to write reports that are free of errors in fundamental writing mechanics (i.e., reports that are characterized by good grammar, punctuation, spelling, and word choice).

ABILITY 3 - INFORMATION AND ELEMENTS: The ability to include all necessary information and elements in reports.

ABILITY 4 - TIMELINESS: The ability to write acceptable reports in a timely manner.

ABILITY 1--ORGANIZATION AND NARRATIVE: The ability to write clear and organized narratives in reports.

Using the scale below, decide which scale value best describes the reports written by the trainee you are evaluating. Record your evaluation above on the line to the left of the definition.

ORGANIZATION AND NARRATIVE

- 5 = <u>Excellent:</u> Reports are fluent, well developed, and well organized. They show clear command of language and are clearly and logically presented. No, or very little, editing is required to improve the narrative.
- 4 = <u>Good</u>: Despite occasional faults, reports are generally well written and well organized. They are less fluent and less detailed than an excellent report, but demonstrate greater facility than an adequate report. Reports at this level require little, if any, editing to improve the clarity of the narrative.
- 3 = <u>Adequate:</u> Reports demonstrate basic writing competence though they may contain problems in sentence structure or diction, or have limited development. Occasionally, some revision is required to ensure proper interpretation. This represents the minimum acceptable level of performance.
- 2 = <u>Poor</u>: Reports often require further development, lack detail and specificity, or are poorly organized. Reports at this level typically require extensive revision and rewriting.
- 1 = <u>Very Poor:</u> Reports lack coherence and/or adequate development. Reports at this level are not suitable for revision.

<u>ABILITY 2--WRITING MECHANICS</u>: The ability to write reports that are free of errors in fundamental writing mechanics (i.e., reports that are characterized by good grammar, punctuation, spelling and word choice).

Using the scale below, decide which scale value best describes the reports written by the trainee you are evaluating. Record your evaluation on the appropriate line above.

WRITING MECHANICS

- 5 = **Excellent:** Reports show a clear command of the language and generally contain very few, in any, errors in grammar, punctuation spelling, or word choice. Reports require no, or very little, editing to correct technical writing faults.
- 4 = <u>Good</u>: Despite occasional technical writing faults, reports are generally well written and require limited editing.
- 3 = <u>Adequate:</u> Reports demonstrate basic competence but usually contain some errors in grammar, punctuation, spelling or word choice. Reports at this level sometimes require revision to ensure proper interpretation. This represents the minimum acceptable level of performance.
- 2 = <u>Poor:</u> Reports are marred by frequent errors in sentence structure, punctuation, spelling, or word choice. Problems with mechanics make editing for correctness extremely difficult.
- 1 = <u>Very Poor:</u> Reports contain too many technical errors to correct. Reports at this level are not suitable for revision.

<u>ABILITY 4--TIMELINESS</u>: The ability to write acceptable reports in a timely manner.

Using the scale below, decide which scale value best describes the reports written by the trainee you are evaluating. Record your evaluation on the appropriate line above.

TIMELINESS

- 5 = <u>Excellent</u>: Always produces accurate, well written reports in less time than is typical.
- 4 = <u>Good</u>: Often produces accurate, well written reports in less time than is typical.
- 3 = <u>Adequate</u>: Usually produces accurate, well written reports within a reasonable period of time.
- 2 = <u>Poor</u>: Often requires an excessive amount of time to produce an accurate, well written report.
- 1 = <u>Very Poor</u>: Always requires an excessive amount of time to produce an accurate, well written report.

ABILITY 3--INFORMATION AND ELEMENTS: The ability to include all necessary information and elements in reports.

Using the scale below, decide which scale value best describes the reports written by the trainee you are evaluating. Record your evaluation on the appropriate line above.

INFORMATION AND ELEMENTS

- 5 = <u>Excellent</u>: Reports contain all essential and relevant information as well as all elements. Information and elements are clearly and logically presented. No, or very little, editing is required.
- 4 = <u>Good</u>: All essential information and elements are present, but reports may contain minor omissions of relevant information. Information and elements are not as clearly and logically presented as in excellent reports. Little editing is required of reports at this level.
- 3 = <u>Adequate</u>: Essential information and elements are present, but there may be omissions of relevant information and the elements may not be clearly presented. Parts of the reports may have to be rewritten to ensure proper interpretation. Information and elements are presented just well enough to satisfy minimum requirements.
- 2 = <u>**Poor**</u>: Essential information and elements are omitted. Reports at this level typically require extensive revision and rewriting.
- 1 = <u>Very Poor</u>: Much necessary information and many elements are omitted. Reports at this level are not suitable for revision.

Section IV: OVERALL JOB PERFORMANCE

Finally, you are asked to evaluate the trainee's overall job performance using the 5-point scale below.

What is the trainee's overall level of job performance?

ERY POOR	POOR	ADEQUATE	GOOD	EXCELLENT
1	2	3	4	5
Far Below Job Demands	Often Unable to Meet	Just Meets Job Demands	Often Exceeds Job	Far Exceeds Job Demands

N = Not Observed/Unable to rate

Overall Job Performance: Includes all of the different aspects of job performance that you have reviewed today.



This concludes the performance appraisal. Thank you for your participation. Please forward this booklet to your department coordinator.



APPENDIX F

FIELD TRAINING SUCCESS/FAILURE

DATA COLLECTION INSTRUCTIONS AND CODING SHEET



Commission on Peace Officer Standards & Training

FIELD TRAINING SUCCESS DATA

<u>Materials</u>

1.

- 1. Instructions for coding data
- 2. Field Training Success Coding Sheets

Special coding sheets are enclosed to be used to record information about selected officers' success or failure in completing **field training**. The coding sheets contain the names and Social Security Numbers of officers selected for the POST Physical Test and Reading & Writing Test study. For each officer listed, you are to use the below **Coding Instructions** to provide the requested information.

The completed coding sheets are to be returned to POST in accordance with the project deadline (see your agency coordinator for the due date).

If you have any questions, call John Weiner at (916) 739-3886. Send completed forms to: Commission on POST, 1601 Alhambra Blvd., Sacramento CA 95816, Attn: John Weiner.

CODING INSTRUCTIONS

The below items are to be coded for each listed officer, as follows:

BASIC COMP DATE: Date completed basic training, month/day/year (MM/DD/YY)

2. <u>EMP STATUS</u>: current employment status

A = Active employee

- S = Separated
- O = Other (e.g., suspended, disability leave, etc.)
- 3. <u>SEP DATE</u>: Date separated from department (month/day/year) if not currently employed.
- 4. <u>FIELD TRN OUTCOME(S)</u>: Success or failure in completing field training. See codes on reverse side. <u>Note</u>: multiple reasons for failure may be coded (up to 3 reasons).
- 5. <u>FT COMP DATE</u>: Date completed field training (month/day/ year). Leave blank if officer separated during field training.

(over)
CODES FOR FIELD TRAINING OUTCOMES

* NO OUTCOME *

i = in-progress

* COMPLETED *

C1 = Completed in normal time

- C2 = Completed -- required extra time (remediation)
- C3 = Completed -- time required unknown

* FAILED TO COMPLETE *

化合合物 化二氟乙酸 氯磺酸乙酸 植物的

RESIGNED (VOLUNTARY)

R1 = Overall job performance was satisfactory.

Unsatisfactory performance due to ...

R2 = ... inadequate physical ability .

- R3 = ... inadequate report writing skills.
- R4 = ... inadequate analytical skills.
- R5 = ... inadequate job knowledge, skills or abilities other than physical/report writing/analytical.
- R6 = ... other reasons (e.g., attitude, motivation, work habits).

R7 = Performance level unknown.

TERMINATED (INVOLUNTARY)

T1 = Overall job performance was satisfactory.

Unsatisfactory performance due to

- T2 = ... inadequate physical ability .
- T3 = ... inadequate report writing skills.
- T4 = ... inadequate analytical skills.
- T5 = ... inadequate job knowledge, skills or abilities other than physical/report writing/analytical.
- T6 = ... other reasons (e.g., attitude, motivation, work habits).

T7 = Performance level unknown.

FAILED BUT CONTINUED IN NON-PATROL ASSIGNMENT

Unsatisfactory performance due to ...

- F2 = ... inadequate physical ability .
- F3 = ... inadequate report writing skills. F4 = ... inadequate analytical skills.
- F4 = ... inadequate analytical skills.
- F5 = ... inadequate job knowledge, skills or abilities other
 - than physical/report writing/analytical.
- F6 = ... other reasons (e.g., attitude, motivation, work habits).

OTHER

- O1 = Iniurv
- O2 = Other (retired, transferred, etc.). A subsect to a set the control of the set

1997,2010

Commission on Peace Officer Standards and Training

>>> CODING SHEET #1: Field Training Success Data <<<

(2)

(3)

(4)

(5)

(1)

<u>Officer Name</u>	<u>ssn</u>	Basic Comp <u>Date</u>	Emp <u>Status</u>	Sep Date	Field Trn Outcomes* (A) (B) (C)	FT Comp <u>Date</u>
				_1_1_		//
		11		1.1		
			· · · · · ·			
		///	· · · · · · · · ·			//
				//	· · · · · · · · · · · · · · · · · · ·	//
		//		_/_/		·//
		/		_/_/_	· · · · · · · · · · · · · · · · · · ·	_/_/_
	**	//		_!_!		_/_/_
		/_/_		_//		_/_/_
	•	/_/_		_1_1_		
	na territoria di territoria. =	//	· · · · ·	_/_/_		11
		1 1 1 1		1 1		
				· · · · · · · · · · · · · · · · · · ·		
				''		''
						//
				//_	<u> </u>	_/_/_
		/_/_		_/_/		· · · · · · · · · · · · · · · · · · ·
		/_/_		_/_/		_/_/_
		//	<u> </u>	_/_/_	· · · · · · · · · · · · · · · · · · ·	_/_/_
		//	· · · ·	_/_/_		//
		//		_1_1_	مىلىك رىمىيە مىلىم	

*Note: Multiple reasons for failure may be coded

a part me persets satisfies as interrest

The second graviant areas of a tates of a

÷		$(x_1, x_2) = \sum_{i=1}^{n} \frac{(x_1^{i} - x_2^{i})}{(x_1^{i} - x_2^{i})}$	Carlor St.		
	and the second	ci de	i de la compañía de la		
	and a start of the second s	$\frac{1}{2} = \frac{1}{2} $	$\{\hat{Q}_{i}\}_{i=1,\dots,N} = \{\hat{Q}_{i}\}_{i=1,\dots,N}$	$\sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{j=1}^{n-1} \sum_{j=1}^{n-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} $	
					د بر ۲۰۰۰ میلید میلید محکوم

(1) A second s second s second secon second sec

a formar and the second second second for the second second second second second second second second second s

to an internet and the second second

APPENDIX G

MULTIPLE REGRESSION ANALYSES READING & WRITING TESTS AND ESSAY TEST

PREDICTING BASIC TRAINING AND FIELD TRAINING PERFORMANCE



1.1.1

i ne nen ne nen iner a leest

(multreg.sas) Multple Regression Analysis - predicting academy ratings of writing ability Read/Write Tests & Essay Test 07:46 Monday, May 23, 1994

	1 C		
16-61	~ 1 *	NODET	
MOG	ет:	TODELI	
			1 - 1

Dependent Variable: AVGACAD Avg Academy Rating on Report Writing

Analysis	of Variance

Source	DF	Sum Squa:	of M res Squ	lean Iare H	7 Value	Prob>F
Model	6	24.48	788 4.08	131	13.321	0.0001
Error	406	124.38	735 0.30	637		
C Total	412	148.87	523		n an an Arian An Arian	
Root MSE		0.55351	R-square	0.164	15	
Dep Mean		3.67343	Adj R-sq	0.152	21	

3.67343 15.06794

c.v.

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter=0	Prob > T
INTERCEP	1	1.951847	0.20664560	9.445	0.0001
SPELL	1	0.003921	0.00218303	1.796	0.0732
CLARITY	1	0.001028	0.00240164	0.428	0.6687
VOCAB	1	0.001648	0.00236832	0.696	0.4869
MC READ	ĩ	0.006029	0.00215365	2.800	0.0054
CLOZE	ī	0.003983	0.00287993	1.383	0.1674
ESSAYTOT	ī	0.059892	0.01950146	3.071	0.0023
		Standardized	Variable		
Variable	DF	Estimate	Label		
INTERCEP	1	0.0000000	Intercept		
SPELL	1	0.09394153	Spelling (% so	core)	
CLARITY	1	0.02300402	Clarity (% sco	ore)	
VOCAB	ī	0.03849986	Vocabulary (%	score)	
MC READ	ī	0.15794267	Multiple Choid	ce Reading (% s	score)

MC READ	_1	0.15794267	Multiple Choice Reading (8	score)
CLOZE	1	0.08444425	Cloze Test (% score)	
	-			

ESSAYTOT 1 0.16210702 Essay Total Score

(excludes cases tested less than 1 day before academy end date)

加大的风险的

an high the same and the second

Table H-1

Empirical Expectancy Table Reading & Writing Test Scores Predicting Academy Graduation vs. Failure/Withdrawal for Academic Reasons Total Sample -- 1987 Study

Read/Write Cut Score	% Gra	% Gain vs. base rate	
	Achieve cut score		
55	98.7%	92.0%***	4.1%
50	97.2%	90.9%***	2.6%
45	97.1%	87.9%***	2.5%
40	96.2%	85.6%***	1.5%
35	95.7%	82.6%***	1.0%
	Base rate = 94.	7% (N=1271)	

Note: Percent gain=((percent graduate and achieve cut score/base rate percent)-1)*100. Significant differences (Chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score (not shown) denoted as follows: ***p<.0001 (one-tailed).

¹Academy success/failure index: Graduated=1; Failed or withdrew for academic reasons=0; remaining cases were excluded from the analyses.

117 (소설) Table H-2

Empirical Expectancy Table Reading & Writing Test Scores Predicting Academy Graduation vs. Failure/Withdrawal for Academic Reasons by Race/Ethnicity and Gender -- 1987 Study

R/W Cut Score	Asian	Black	Hispanic	White	Male	Female
\$ 5	100%	95.0%	95.3%	99.3% ***	99.1% ***	96.5%**
	(9.1%)	(5.9%)	(3.3%)	(2.9%)	(3.5%)	(7.6%)
50	100%*	95.0%	93.3%	97.9%***	98.6%***	90.3%
	(9.1%)	(5.9%)	(1.2%)	(1.4%)	(3.0%)	(0.7%)
45	100%***	93.3%	95.2%	97.8%***	98.5%***	90.6%
	(9.1%)	(4.0%)	(3.2%)	(1.3%)	(2.8%)	(1.0%)
40	95.0%	90.0%	95.5%**	97.3%***	97.4%***	90.8%
	(3.6%)	(0.3%)	(3.5%)	(0.8%)	(1.7%)	(1.2%)
35		91.1% (1.6%)	94.8%** (2.7%)	96.8% * (0.3%)	96.9%*** (1.2%)	89.9% (0.3%)
Base rate	91.7%	89.7%	92.3%	96.5%	95.7%	89.7%
Ň	48	107	181	886	1058	194

Percent Graduating¹ and Achieving Cut Score (% Gain relative to base rate shown in parentheses)

Note: Percent gain=((percent graduate and achieve cut score/base rate percent)-1)*100. Significant differences (Chi-square or Fisher's exact test) between percent achieve cut score vs. percent below cut score (not shown) denoted as follows: *p<.05, **p<.01, ***p<.001 (one-tailed).

¹Academy success/failure index: Graduated=1; Failed or withdrew for academic reasons=0; remaining cases were excluded from the analyses.