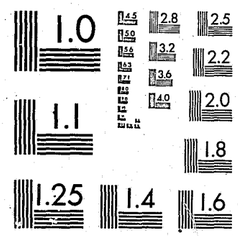

MF-1

VALIDATION REPORT

VALIDATION OF ENTRY-LEVEL POLICE OFFICER EMPLOYMENT TESTS

EMPLOYMENT STANDARDS SECTION
MICHIGAN LAW ENFORCEMENT OFFICERS TRAINING COUNCIL
7426 N. CANAL ROAD, LANSING, MI 48913, PHONE: (517) 322-1946

VALIDATION OF ENTRY-LEVEL POLICE OFFICER EMPLOYMENT TESTS

Prepared for the:

EMPLOYMENT STANDARDS SECTION

MICHIGAN LAW ENFORCEMENT OFFICERS TRAINING COUNCIL

WOLLACK & ASSOCIATES
A Psychological Corporation

June, 1981

This project is supported by Grant No. 12266-7A79 awarded to the Michigan Law Enforcement Officers Training Council by the Michigan Office of Criminal Justice and the Law Enforcement Assistance Administration, U.S. Department of Justice. Points of view presented or information provided in this document are those of the authors and do not necessarily represent the official position of the Michigan Office of Criminal Justice or the U.S. Department of Justice.

Printed by the State of Michigan.

****

For more information regarding this project, please contact:

EMPLOYMENT STANDARDS SECTION
Michigan Law Enforcement Officers Training Council
Department of State Police
7426 N. Canal Road
Lansing, MI 48913
PH: (517) 322-1946

****

August 1981

Dear Colleague:

This Report is one in a series of reports documenting the Training Council's standards validation effort. The validation of job-related selection and training standards for entry-level Michigan police officers is the highest priority of the Training Council. The work products of this project will ensure the fair and equitable selection of police officer candidates. The employment tests will become part of the state's Minimum Employment Standards and the recruit training curriculum will be validated and updated.

This effort would not have been possible without the cooperation and contributions of Michigan's law enforcement agencies and management and labor organizations. The many hours of participation in the validation effort will guarantee that the standards are directly linked to the police officer job.

On behalf of the Training Council, I want to thank the Michigan law enforcement community for their contribution to this significant step forward for our profession.

Sincerely,

William Lucas

William Lucas
Chairman

PREFACE

This Validation Report is an abridged edition of the Technical Report by the same name. This abridgement was done in the name of economy and will allow wider distribution of the essential components of this important document. This Report contains only the body from the Technical Report. What has been deleted are the appendices which are essentially supporting documentation and are not of use to the typical reader.

Copies of the appendices are available for inspection at the Employment Standards Section, Michigan Law Enforcement Officers Training Council, 7426 N. Canal Road, Lansing, MI 48913, (517) 322-1946.

TABLE OF CONTENTS

I. INTRODUCTION 1

II. BASIC LITERACY EXAMINATION--UNIFORM GUIDELINES: CONTENT VALIDITY 2

III. READING COMPREHENSION EXAMINATION 6

- Purpose of Readability Analysis 6
- The Collection of Source Materials 9
- Linkage Analysis 9
- Result of Readability Analysis 10
- Examination Development 14
- Alternate Showing Analysis 16

IV. WRITING SKILLS EXAMINATION 23

- The Report Writing Survey Form 23
- Examination Development 28
- Alternate Showing Analysis 30

V. USE OF THE BASIC LITERACY EXAMINATION 33

VI. BASIC PHYSICAL SKILLS TEST BATTERY--UNIFORM GUIDELINES: CONTENT VALIDITY 35

VII. PHYSICAL SKILLS TEST 39

- Sample Selection 40
- Normative Testing 43
- Selection of Final Test Battery 44
- Evidence of Validity 45
- Recommended Computational Procedure 49
- Comparison of Incumbent Officers by Sex 51
- Administration of the Modified Battery 54

VIII. USE OF THE PHYSICAL SKILLS EXAMINATION 59

- A Pass-Fail Examination 59
- Method of Test Scoring 60

LIST OF TABLES

No.	Title	Page
1	Readability Analysis (N = 342)	10
2	Readability Analysis of Police Training Materials	14
3	Item Characteristics of Two Reading Comprehension Examinations	15
4	Racial Composition of Normative Samples	17
5	Normative Sample for Basic Literacy Examination	18
6	Comparison of Reading Comprehension Test Performance by Racial Groups	21
7	Departments Providing Data Base for Report Writing Survey	25
8	Description of Field Report Usage by Patrol Officer Job Tasks (N = 35)	26
9	Relative Importance of Report Writing Requirements (N = 32)	27
10	Relative Weights for the Five Writing Skills Items	29
11	Item Characteristics of Two Writing Skills Examinations	30
12	Comparison of Writing Skills Test Performance by Racial Groups	31
13	Preliminary Physical Skills Test Battery	39
14	Agencies That Participated in Physical Testing	40
15	Police Officer Physical Skills Test Sample	42
16	Comparison of Work Sample Tests to Job Analysis Findings	46
17	Correlational Data: Six-Event Battery	47
18	Weighting of Examination Events	49
19	Comparison of Incumbent Officers by Sex on Physical Skills Test Battery	51
20	Comparison of Acceptance Rates by Sex for Incumbent Police Officer Sample	52
21	Descriptive Data: Six-Event Battery (Incumbents)	53
22	Descriptive Data: Revised Test Battery (Students)	55
23	Percent of Examinees Who Failed to Complete Events	57

I. INTRODUCTION

In April, 1980, Wollack & Associates, A Psychological Corporation, contracted with the Michigan Law Enforcement Officers Training Council to develop and validate entry-level patrol officer testing procedures. This report summarizes the research studies performed by Wollack & Associates in the development of a basic literacy examination and a basic physical skills test battery.

The purpose of the basic literacy examination is to assess job applicants' reading comprehension and report writing skills to determine his or her suitability for employment as a police officer. These important job skills were determined by previous research studies to be essential requirements for training and job performance by entry-level police officers. The basic physical skills test battery is intended to measure those physically-demanding aspects of police officer job performance which are deemed to be essential and may reasonably be expected of job applicants lacking police training. The intent of this research undertaking is to provide job-related selection instruments which may be used as the basis of a pre-employment testing program for selecting entry-level personnel into all law enforcement agencies with patrol responsibilities. Such a testing program is essential for purposes of identifying qualified, competent job applicants for police positions in a manner which is non-discriminatory. The validation studies reported herein sought to comply with the Uniform Guidelines on Employee Selection Procedures (1978) which have been published by the federal compliance agencies.

Dr. Stephen Wollack was responsible for the technical supervision of this research project. Dr. Merle Foss, an exercise physiologist and subcontractor, assisted greatly in the development of the physical skills test battery. Mr. Eric Sayenga, a subcontractor, provided assistance in the data analysis phase of this project. Special thanks are due the Employment Standards Section of the MLEOTC for their efforts in overseeing the administrative aspects of the research study. In particular, Mr. Patrick Judge, Mr. William Nash, and Mr. Dale Rothenberger should be acknowledged for the highly professional, competent manner in which they oversaw and coordinated this important research study. Finally, gratitude should be expressed to the many police agencies throughout the State of Michigan which participated in this project and helped to fashion the final work product.

II. BASIC LITERACY EXAMINATION
UNIFORM GUIDELINES: CONTENT VALIDITY

The following index describes the citations, the corresponding requirements, and a listing of reference pages indicating the appropriate sections of this validation report which deal explicitly with the content validation requirements of the Uniform Guidelines on Employee Selection Procedures adopted by the Equal Employment Opportunity Commission, the U.S. Civil Service Commission, the U.S. Department of Labor, and the U.S. Department of Justice (August 25, 1978). This listing includes all reporting requirements which have been designated as being essential for documenting the content validity of employment tests.

Citation	Requirement	Reference
15 C (1)	Dates and location(s) of the job analysis should be shown.	pp. 6-9 pp. 23-25 Appendix I
15 C (3)	A description of the method used to analyze the job should be provided.	pp. 6-9 pp. 23-24 Appendix G
15 C (3)	The work behavior(s), the associated tasks and, if the behavior results in a work product, the work products should be completely described.	See: Statewide Job Analysis of the Police Patrol Officer Position, 1979, Appendices S-1, L-37
15 C (3)	Measures of criticality and/or importance of the work behavior(s) and the method determining these measures should be provided.	Ibid

Citation	Requirement	Reference
15 C (3)	Where the job analysis also identified the knowledges, skills, and abilities used in work behavior(s), an operational definition for each knowledge in terms of a body of learned information and for each skill and ability in terms of observable behaviors and outcomes, and the relationship between each knowledge, skill, or ability and each work behavior, as well as the method used to determine this relationship, should be provided.	pp. 9-14 pp. 23-28
15 C (4)	Selection procedures, including those constructed by or for the user, specific training requirements, composites of selection procedures, and any other procedures supported by content validity, should be completely and explicitly described or attached.	pp. 14-16 pp. 28-30
15 C (4)	If commercially available selection procedures are used, they should be described by title, form and publisher.	Not applicable
15 C (4)	The behaviors measured or sampled by the selection procedure should be explicitly described.	pp. 14-16 pp. 28-30
15 C (4)	Where the selection procedure purports to measure a knowledge, skill, or ability, evidence that the selection procedure measures and is a representative sample of the knowledge, skill, or ability should be provided.	pp. 10-15 pp. 27-29
15 C (5)	The evidence demonstrating that the selection procedure is a representative sample of the work behavior(s), or a representative sample of a knowledge, skill, or ability as used as part of a work behavior and necessary for that behavior should be provided.	pp. 10-16 pp. 23-30

Citation	Requirement	Reference
15 C (5)	The user should identify the work behavior(s) which each item or part of the selection procedure is intended to sample or measure.	p. 6 p. 27
15 C (5)	Where the selection procedure purports to sample a work behavior or to provide a sample of a work product, a comparison should be provided of the manner, setting, and the level of complexity of the selection procedure with those of the work situation.	pp. 10-15 pp. 23-29
15 C (6)	The alternative selection procedures investigated and available evidence of their impact should be identified.	pp. 16-22 pp. 30-32
15 C (6)	The scope, method, and findings of the investigation, and the conclusions reached in light of the findings, should be fully described.	pp. 6-32
15 C (7)	The methods considered for use of the selection procedure . . . and available evidence of their impact should be described.	pp. 21-22 pp. 31-34
15 C (7)	This description should include the rationale for choosing the method for operational use, and the evidence of the validity and utility of the procedure as it is to be used.	pp. 33-34
15 C (7)	The purpose for which the procedure is to be used (e.g., hiring, transfer, promotion) should be described.	p. 1
15 C (7)	If the selection procedure is used with a cutoff score, the user should describe the way in which normal expectations of proficiency within the work force were determined and the way in which the cutoff score was determined.	pp. 20-22 pp. 30-32

Citation	Requirement	Reference
15 C (7)	In addition, if the selection procedure is to be used for ranking, the user should specify the evidence showing that a higher score on the selection procedure is likely to result in better job performance.	Not applicable
15 C (8)	The name, mailing address, and telephone number of the person who may be contacted for further information about the validity study should be provided.	William C. Nash MLEOTC 7426 N. Canal Rd. Lansing, MI 48913 (517)-322-1946

III. READING COMPREHENSION EXAMINATION

In a previously reported research study performed on behalf of the Michigan Law Enforcement Officers Training Council, Personnel Research Consultants identified a variety of reading requirements which are common among traditional and specialized agencies (see: Statewide Job Analysis of the Police Patrol Officer Position, 1979). Appendix S-1 of that report provides a comprehensive listing of reading subject matter across the 12 agency types which questionnaire respondents identified as being relevant to their jobs. Clearly, these data demonstrate that reading comprehension is a fundamental job skill which cuts across all varieties of law enforcement agencies. Moreover, the ability to read with comprehension is an essential training requirement which is common to all police personnel. Numerous previously conducted research studies of police officer job requirements have consistently shown reading skills to be a basic requirement in both training and one's day-to-day job behavior.

Purpose of Readability Analysis

This report summarizes the readability analysis conducted on a large sample of reading subject matter utilized by police officers in training and in their jobs. The objective of this research was to provide the necessary data supporting the job-relatedness of a reading comprehension examination used as part of a screening process for evaluating entry-level law enforcement personnel.

The readability analysis described in this report constitutes the basis for the definition of the test's "content domain". An analysis of the degree of content validity of an employment test should be based upon the degree to which the content domain of the test matches the job content domain. To this end, a readability analysis was undertaken to determine the reading difficulty level of subject matter which must be read by police officers.

The index used in this research for determining reading difficulty is the Farr-Jenkins-Paterson Index.

The procedure used in calculating the Farr-Jenkins-Paterson Index is as follows:

1. Select in a systematic manner samples of 100 words;
2. Divide the number of words by the number of sentences to determine average sentence length (SL);
3. Count the number of words having one syllable and divide by the total number of words (in the 100 word sample) to determine the proportion of one syllable words (nosw);
4. Calculate the Farr-Jenkins-Paterson Index by inserting the above values into the following equation, 1.599 (nosw)--1.015 (SL)--31.517.

Using this formula, the higher the index obtained the easier the passage is to read and understand. The following gives index ranges for different difficulty categories.

80 - 89	Easy
70 - 79	Fairly easy
60 - 69	Standard
51 - 59	Fairly difficult
30 - 50	Difficult
0 - 29	Very difficult

There are a large number of other readability formulas that may be utilized in determining the comprehensibility of written passages. One researcher (Klare, 1963) has systematically analyzed 31 different readability formulas. To choose from among this array, it was necessary to establish three criteria. The criteria for selecting a formula were the following:

Accuracy. The accuracy of readability formulas is generally determined by reference to a set of standardized criterion passages, or by reference to agreement with other formulas which have been previously evaluated for accuracy. Powers, et al. (1958) utilized a regression technique to demonstrate a high degree of agreement between the Flesch, The Dale-Chale, the Farr-Jenkins-Paterson, and the Gunning (FOG) readability formulas. The authors conclude that the differences in accuracy between the indices analyzed are small enough to be of little practical significance. An analysis of the various methodologies

indicates that the Farr-Jenkins and FOG indices yield readability values that are quite similar to other readability measures commonly used, and this high degree of correspondence is consistent throughout all levels of difficulty. Another important aspect of index accuracy is the reliability of the measurements. In calculating a large number of reading index scores, the simplest technique is likely to be the most reliable. By this standard alone, the Farr-Jenkins-Paterson Index is preferable to most other reading indices because of its simplified counting and calculation procedures.

Convenience. In undertaking readability analysis on a large number of passages, it is desirable to utilize an index that will minimize the amount of time required per passage. As prevously noted, the Farr-Jenkins-Paterson Index is clearly superior on this criterion. With minimum training and experience, an analyst can quickly and accurately conduct readability analysis on several passages with no significant loss of accuracy due to fatigue or distraction. The convenience of this formula also allows researchers to monitor thoroughly all readability analysis by independently checking a significant portion of passages sampled.

Suitability. The final criterion for selecting the Farr-Jenkins-Paterson Index is its appropriateness for use with the material under investigation. Of the 31 formulas reviewed by Klare (1963), nearly one-third of these formulas were not suitable for use with material to be read by adults. Hence the choice of formulas was somewhat limited.

For the purpose of this research, the Farr-Jenkins-Paterson Index is significant because it provides a common standard to evaluate the difficulty level of job materials and the difficulty level of test content. What is important is that individuals be able to read and comprehend material that is of the same difficulty as material they will encounter as police officers. Thus, we have utilized this index in analyzing the reading difficulty of a large sample of reading materials encountered by police officers. Further, we have utilized the same index in analyzing the reading difficulty of test passages specifically designed to measure reading comprehension of police job materials. Comparison of reading difficulty levels for job materials and for test materials is appropriate and meaningful only when the same index is used on both sets of materials. In this manner, the

degree of correspondence (i.e., content validity) between the test and job requirements may be clearly demonstrated.

All written materials analyzed for reading difficulty levels were provided by the MLEOTC staff. The reading subject matter furnished to Wollack & Associates included materials which officers were required to read and understand. From these pages, a total of 342 samples of 100 words (or more) were analyzed using the Farr-Jenkins-Paterson Index.

The Collection of Source Materials

On april 16, 1980, letters were sent to 21 MLEOTC training academies, both basic and pre-employment programs, requesting that they submit bibliographies of materials required to be read by trainees. They were also requested to send all indicated reading materials (see Appendix A). All 21 academies responded with the requested information. A listing of those training facilities is provided as Appendix B. Moreover, a bibliography and associated reading materials were submitted, as well, by the Department of Natural Resources. The reading subject matter which became the basis for the readability analysis included: case law; statutes, court decisions, the Criminal Law and Procedures Manual; first aid manuals; the Michigan Liquor Control Act; the Michigan Vehicle Code, and training bulletins.

Linkage Analysis

Concurrent with the development of the entry-level examinations, the MLEOTC contracted with a consulting firm, Psychological Services Inc., to revise the current training curriculum in order to conform with applicable standards of job-relatedness. This, of course, is a highly meaningful and important undertaking, because it helps to establish the content validity of training requirements. As an adjunct to this undertaking, a survey was performed to establish the relationship between the bibliography subject matter and the relevant job behaviors. Specifically, each reading source material was related to the appropriate Training Task Group(s). This linkage was performed by the MLEOTC staff in accordance with instructions provided by Wollack & Associates in the form of a correspondence dated April 1, 1980 (see Appendix C). The reading materials to be analyzed were to be related to the 34 core Training Task Groups (including non-core tasks for Airport Police, Parks & Wildlife, and

Railroad Police). In all, there were 308 such core tasks, as well as an additional 24 non-core tasks which were included, for a total of 332. The reading materials were reviewed for relevance to the task groupings as well as the underlying task statements by the MLEOTC staff. A matrix was prepared containing the source, authors, date, chapters or subparts, as well as the task groups to which the source materials pertained. In this manner, the MLEOTC staff prepared a complete listing of the reading materials linked to the 34 task groups identified in the previous job analysis. The reading linkage documentation, as well as associated correspondences, have been included as Appendix D. Appendix E is additional documentation which shows the various reading source materials and the training facilities which have indicated their use of those materials.

Result of Readability Analysis

Table 1, which follows, provides a listing of the source materials upon which the readability analysis was performed, their author/source, the number of 100 word passages which were sampled, and the mean readability index for that source. The number of passages sampled for any given source was determined by a rule of thumb which took into consideration the form of the materials, the length of the source, etc.

Table 1.
Readability Analysis (N = 342)

Title	Author/Source	Number of Passages	Index
Criminal Investigation and Physical Evidence Handbook	Wisconsin Dept. of Justice Crime Lab Bureau	5	47.4
State of Michigan Official Traffic Accident Report	Dept. of State Police	2	50.0
Police Firearms	National Rifle Association	5	57.6
Standard First Aid and Personal Safety	American Red Cross	10	48.0
Advanced First Aid and Emergency Care	American Red Cross	10	49.3
How to Recognize and Handle Abnormal People	National Association for Mental Health	7	66.7

Table 1 (contd.)

Title	Author/Source	Number of Passages	Index
First Aid for Foreign Body Obstruction of the Airway	American Red Cross	3	44.5
Cardiopulmonary Resuscitation	American Red Cross	5	49.0
Vehicle Damage Severity Scale for Michigan Traffic Accident Investigators	National Safety Council	3	52.2
Criminal Investigation Specific Offenses: Volume Two	International Association of Chiefs of Police	6	51.3
Criminal Investigation Basic Procedures: Volume One	International Association of Chiefs of Police	6	44.8
Traffic & Accident Investigation	Kalamazoo Regional Police Academy	8	48.4
Michigan Criminal Procedure: Cases, Problems & Materials	Ramsdell and McCloskey	10	56.7
Michigan Criminal Law: Cases, Problems & Materials	McCloskey, Ramsdell, and Schroeder	10	49.3
Concealed Weapons and Firearms Laws	Department of State Police	5	41.7
Law Enforcement Officers Manual on the State Liquor Laws and Rules and Regulations of the Commission	State Liquor Control Commission	2	24.3
Types of Driver Licenses Issued by Michigan	Secretary of State	5	40.5
What Every Driver Must Know	Secretary of State	5	63.3
Bomb Threats and Search Techniques	Department of the Treasury	5	57.9
Constitution of the United States	--	5	47.3
Collection and Preservation of Physical Evidence	MLEOTC	5	52.6
Fingerprinting and Palm Printing	MLEOTC	5	52.5

Table 1 (contd.)

Title	Author/Source	Number of Passages	Index
Latent Prints	MLEOTC	5	54.3
Recognition of Physical Evidence	MLEOTC	5	57.6
Radio Communications	MLEOTC	5	51.4
Crime Scene Photography	Department of State Police	5	54.3
Firearms Evidence: Collection Marking, Packing, and Preservation	--	5	43.8
Constitution of the State of Michigan	--	5	35.3
Civil Disorder Control	Department of State Police	5	49.9
Criminalistics	Department of State Police	5	48.6
Michigan's Basic Law of Arrest and Search and Seizure	Lansing Community College	5	38.7
Defensive Tactics Manual	MLEOTC	5	41.6
Search Warrants Manual	Department of State Police	5	42.9
Law Enforcement and Youth	Department of State Police	5	43.4
Handbook of Michigan Criminal Law and Procedures	Lansing Community College	5	38.1
Precision Driving Techniques	Department of State Police	5	62.9
Basic Firearms Training Manual	Department of State Police	5	58.8
Latent Prints	Department of State Police	5	56.3
Michigan Liquor Control Act and Rules	State Liquor Control Commission	5	41.8
Traffic Accident Investigation Manual	J. S. Baker	10	51.5
Fundamentals of Criminal Investigation	C. E. O'Hara	10	53.1
The Patrol Operation	International Association of Chiefs of Police	10	51.7

Table 1 (contd.)

Title	Author/Source	Number of Passages	Index
Michigan Vehicle Code	Secretary of State	10	32.5
Michigan Juvenile Court Procedure Sourcebook	Office of the Court Administrator, Lansing	5	45.3
Policy & Procedures Manuals	State Department of Natural Resources	30	42.3
Public Rights on Michigan Waters	State Department of Natural Resources	5	53.1
Michigan Pleasure Boating	State Department of Natural Resources	5	69.2
Firearms Training Program	State Department of Natural Resources	5	51.7
A Classroom Course in Water Safety and Survival	State Department of Natural Resources	5	45.8
Michigan Off-Road Vehicle Safety Training Course	State Department of Natural Resources	5	62.2
Michigan Snowmobile Safety Manual	State Department of Natural Resources	5	60.6
Michigan Snowmobiler's Safety Handbook	Snowmobile Safety and Certification Committee	5	49.5
State of Michigan--1977--Natural Resources Laws	State Department of Natural Resources	10	32.0
Natural Resources Rules and Orders--1977	State Department of Natural Resources	10	45.2

Based upon the 342 passages of reading subject matter which were analyzed for reading difficulty, the average readability index is 44.1. Table 2, which follows, provides a breakdown of the reading difficulty level of the surveyed materials.

An inspection of Table 2 reveals that the bulk of reading subject matter analyzed falls into the "Difficult--Fairly Difficult" range as defined by the Farr-Jenkins-Paterson Index.

Table 2.
Readability Analysis of Police Training Materials

Reading Ease Category	Index*	Percent of Subject Matter
Very Difficult	0 - 29	.10
Difficult	30 - 50	.42
Fairly Difficult	51 - 59	.29
Standard	60 - 69	.13
Fairly Easy	70 - 79	.04
Easy	80 & above	.02
$\bar{X}$ = 44.1		100%

*Farr-Jenkins-Paterson Index based upon 342 sampled passages.

Examination Development

Two different types of reading comprehension items were generated by Wollack & Associates. One hundred (100) items were prepared utilizing the traditional reading comprehension format which consisted of several paragraphs followed by a number of related questions. A second item pool, also consisting of 100 items, relied upon a somewhat different format. Brief statements consisting of one or two sentences were followed by a few related questions. The latter format, while somewhat more inferential than the former, did not require the applicant to comb through lengthy passages in responding to questions.

The two item pools were administered to criminal justice classes at Ferris State College and Lansing Community College for the purpose of conducting an item analysis. The two item pools were administered to 197 students. At Ferris, the students were both juniors and seniors from a variety of disciplines. The Lansing students were sophomores enrolled in criminal justice classes. The item analysis was conducted using a program at the Michigan State University Testing Office. The intent of this item analysis was to select from each of the two item pools those 50 items which have the best item characteristics and are most representative of the job content domain.

Federal guidelines permit the claim of content validity for well-developed

measures. The best approach to the development of examinations, to assure reliability, is the psychometric technique of item analysis. In the development of a reading comprehension examination, it is desirable to include items which correlate highly with other items in the test. To the extent that a high degree of intercorrelation exists among test items, the examination is said to possess some degree of internal consistency reliability. In addition to assuring a high degree of reliability, item analysis procedures are recognized as highly important for identifying those test items which result in maximum variability and discriminability among the scores of those individuals tested. An additional objective of item analysis is to evaluate the responses of each test item to determine whether the items are of an appropriate level of difficulty, and, further, whether the item alternatives make a meaningful contribution to the examination process (i.e., do they discriminate).

Table 3, which follows, presents the item characteristics of the two reading comprehension examinations.

Table 3.
Item Characteristics of Two Reading Comprehension Examinations

Item Characteristics	Traditional Format	Brief Format
Number of Items	50	50
Mean Readability Index	43.58	43.26
Mean Item Difficulty Index	21	26
Mean Item Discrimination Index	28	34
K-R 20 Reliability	0.83	0.87
Standard Error of Measurement	2.53	2.67

An inspection of Table 3 shows that each of the two examinations was 50 items in length. The mean readability level for the traditional (paragraph) format was 43.58 as compared with a mean readability level of 43.26 for the brief (sentence) format. Both examinations correspond very closely to the readability analysis of the job and training materials. It should be recalled that the mean readability index for those sampled materials was 44.1. Therefore, the mean readability index of both examinations is regarded as being virtually identical with the readability analysis of the 342 sampled passages. The mean item difficulty indices for the traditional and brief format examinations are 21 and 26 respectively.

The difficulty index is the proportion of the total group who got the item wrong. A high index would indicate a difficult item, while a low index would indicate an easy item. The difficulty values obtained for both examinations are considered to be optimal. The mean item discrimination indices for the traditional and brief format examinations are 28 and 34 respectively. The index of discrimination is the difference between the proportion of the upper group who got an item right and the proportion of the lower group who got the item right. This index is dependent upon the difficulty of an item. The obtained discrimination index values are considered to be acceptable. The internal consistency reliability of the examinations was computed utilizing the "Kuder-Richardson 20" reliability formula. The coefficients for the traditional and brief format examinations were 0.83 and 0.87 respectively. Both values are regarded as high levels of reliability given that the examinations are 50 items in length. The standard error of measurement for the traditional format examination was 2.53 as compared with 2.67 for the brief format examination. In the opinion of Wollack & Associates, the item analysis data reported herein are very satisfactory and describe two examinations whose item characteristics are favorable. Wollack & Associates believes that the data strongly support a conclusion of content validity for both examinations.

The reading comprehension examinations are considered to be power tests rather than speeded tests. The time limit to be imposed on this examination is for administrative purposes only. A subsequent administration of these tests was used for the purpose of determining time limits. J. P. Guilford (1956) defines a power test as one which is finished by at least 75% of the examinees.

Alternate Showing Analysis

The objective of this aspect of the research study was to comply with the Federal Uniform Guidelines on Employee Selection (1978) pertaining to a consideration of alternative selection procedures and alternative methods of use. The two reading comprehension examinations previously described were developed specifically for the purpose of addressing this requirement. This analysis was undertaken to determine whether one of the two alternative methodologies for measuring reading comprehension might have a lesser impact upon minorities. Both reading comprehension examinations are considered to be job-related and valid to the same degree. Therefore, the choice of the item format to be adopted for the final examination form would depend upon any differential

showing of adverse impact. Naturally, an examination form having less impact upon protected classes and being equally valid to other alternatives would be preferred.

In order to conduct this study, it was necessary to contact police departments throughout the State to provide samples of minority and non-minority incumbent police officers. In all, a total of 229 police officers were tested, 166 of whom were white and 63 of whom were minority. Table 4 details the racial composition of the two samples.

Table 4.
Racial Composition of Normative Samples

Racial Group	Number
White	166
Black	55
Hispanic	6
Asian	1
American Indian	1
	229

Appendix F is the letter which was sent by the MLEOTC to police departments for the purpose of soliciting participation in this phase of the research. That letter spells out guidelines for participation of incumbent police officers in this test tryout. These guidelines were furnished by Wollack & Associates and incorporated within the MLEOTC letter. Specifically, the sample was limited to officers on patrol and no older than 35 years of age. This was thought to be an important requirement in order to assure comparability between the normative sample and the typical job applicants who are characteristically in their early 20s. Police departments were advised to provide officers whose job performance is considered to be fully satisfactory in all respects. They were cautioned not to send officers whose job performance levels were regarded as being substantially unsatisfactory. On the other hand, they were advised against sending only the very best people as this would not be regarded as a representative sample. Most

importantly, police departments were asked to provide officers who are sincere and motivated to do their best. Table 5 identifies the composition of the normative sample, including agency types, agencies, and the number of white and non-white officers sampled.

Table 5.

Normative Sample for Basic Literacy Examination

Agency Type/Departments	White Officers Tested	Non-White Officers Tested*
Michigan State Police	28	15
Detroit Police Department	22	24
Large Cities, Villages, Townships		
Ann Arbor Police Department	2	2
Flint Police Department	2	2
Grand Rapids Police Department	2	2
Lansing Police Department	2	2
Pontiac Police Department	2	1
Saginaw Police Department	2	1
Southfield Police Department	3	0
Sterling Heights Police Department	4	0
Taylor Police Department	2	0
Warren Police Department	2	0
Subtotal	23	10
Medium Cities, Villages, Townships		
Allen Park Police Department	1	0
Bay City Police Department	3	0
Ecorse Police Department	0	2
Hamtramck Police Department	2	0
Inkster Police Department	0	2
Lincoln Park Police Department	2	0
Marquette Police Department	3	0
Mt. Clemens Police Department	1	1
Portage Police Department	2	0
Port Huron Police Department	2	0
River Rouge Police Department	1	1
Sault Ste. Marie Police Department	2	0
Wayne Police Department	1	0
Wyandotte Police Department	2	0
Subtotal	22	6

Table 5 (contd.)

Agency Type/Departments	White Officers Tested	Non-White Officers Tested*
Small Cities, Villages, Townships		
Bloomfield Hills Police Department	2	0
Burton Police Department	1	0
Clawson Police Department	1	0
Fenton Police Department	1	0
Flushing Police Department	1	0
Grand Blanc Twp. Police Department	2	0
Holly Police Department	1	0
Ishpeming Police Department	1	0
Mt. Morris Twp. Police Department	1	0
Negaunee Police Department	1	0
Northern Michigan University	1	0
Northville Police Department	1	0
Oakland University	1	0
Oxford Police Department	1	0
Walled Lake Police Department	1	0
Western Michigan University	0	2
White Lake Twp. Police Department	2	0
Wixom Police Department	1	0
Subtotal	20	2
Small Sheriff Departments		
Barry County	1	0
Chippewa County	1	0
Jackson County	2	0
Marquette County	1	0
Midland County	1	0
Montcalm County	1	0
Subtotal	7	0
Large Sheriff Departments		
Allegan County	2	0
Bay County	3	0
Eaton County	3	0
Ingham County	3	0
Kent County	4	0
Lapeer County	2	0
Macomb County	5	0
Monroe County	5	0
Oakland County	2	2
Ottawa County	2	0
St. Clair County	3	0
Wayne County	2	3
Subtotal	36	5

Table 5 (contd.)

Agency Type/Departments	White Officers Tested	Non-White Officers Tested*
Department of Natural Resources	3	2
Parks		
Kalamazoo Co. Parks & Recreation	1	0
Airports		
Tri-County Airport	1	0
Railroads		
Conrail	2	0
Chessie	1	0
Subtotals	8	2

*One non-white officer was unable to participate in the testing. Therefore, the non-white sample was 63 rather than 64.

Section 5H of the Uniform Guidelines specifies: "Where cutoff scores are used, they should normally be set so as to be reasonable and consistent with normal expectations of acceptable proficiency within the work force." Of course, the matter of what is reasonable and consistent with such expectations of job proficiency is intrinsically judgmental. The use of a normative sample comprised of job incumbents is intended to provide a basis for making this judgment. It should be shown that any cutoff score adopted for use would have the effect of passing the vast majority of incumbent officers in the normative sample. Therefore, the 15th percentile was selected for the purpose of establishing the cutoff score. In other words, a cutoff score would be established on each form of the reading comprehension examinations such that an overall acceptance rate of 85% would be produced. Wollack & Associates believes that the adoption of this low cutoff score is certainly compatible with the language of the Guidelines calling for reasonableness. The effect of such a cutoff is to eliminate only those individuals whose reading comprehension skills are extremely low relative to the incumbent officer sample.

Table 6 provides a comparison of test performance on the reading comprehension examinations for white and non-white officers.

Table 6.

Comparison of Reading Comprehension Test Performance by Racial Groups

Test	Statistics	Whites (N = 166)	Non-Whites (N = 63)
FORM A			
	Mean	41.71	37.40
	Standard Deviation	4.42	5.50
	Acceptance Rates	.92	.75
FORM B			
	Mean	44.36	40.43
	Standard Deviation	3.77	5.58
	Acceptance Rates	.89	.75

Form A of the reading comprehension examination consists of the brief format items previously described. Form B consists of the traditional format items. A comparison of the test descriptive statistics is presented above in Table 6. The mean test scores of whites and non-whites on the two examination forms are actually quite close, with a difference of approximately one standard deviation in favor of the white officers. Using the 85% acceptance rate rule of thumb, a raw score cutoff of 36 was adopted for Form A and a raw score cutoff of 39 was adopted for Form B. It should be noted that the use of these raw score cutoffs produced an overall acceptance rate roughly equal to 85%. There is little to choose between the two examination forms with respect to the relative performance of the two racial groups. For Form A, the white officers passed at a 92% rate, while the non-whites passed at a 75% rate. The minority acceptance rate as a percentage of the white acceptance rate is approximately 82%. On Form B, the white officers passed at an 89% rate, while the minority officers passed at a 75% rate. For Form B, the minority acceptance rate as a percentage of the white acceptance rate is 84%. Therefore, the minority officers fared somewhat better in relation to the white officers on Form B of the examination, though the difference between the two forms is regarded

as being small. It should be noted that neither form of the reading comprehension examination produced an adverse impact against the minority officers under the 80% rule recognized by the Federal Uniform Guidelines. Because the minority acceptance rate relative to the white acceptance rate was somewhat better on Form B of the examination (traditional format), it is recommended that Form B be adopted for use within the final form of the examination. A raw score cutoff of 39 has been recommended.

IV. WRITING SKILLS EXAMINATION

The statewide job analysis by Personnel Research Consultants identified 16 core tasks involving writing skills by police officers. These findings are consistent with numerous research studies previously conducted of police officer job requirements. The preparation of reports by police personnel is a frequent and important job requirement. It is, therefore, justified to undertake the development of a writing skills examination which is job-related to the requirements of law enforcement officers.

The Report Writing Survey Form

In order to obtain a more precise description of police patrol officer report writing requirements, a survey questionnaire was prepared which is entitled: Report Writing Survey Form for Law Enforcement Officers (see Appendix G). This survey form consisted of three subsections. Part I requested the names of report forms utilized by patrol officers in the surveyed police departments for each of the 16 core tasks requiring writing skills which were identified in the Kohls and Berner job analysis study. Respondents were asked to identify by name, for each core writing-related task, the form or forms which patrol officers must complete in order to carry out the particular job behavior identified. Additionally, respondents were asked to indicate the frequency of usage (per month) of each identified form.

Part II of the survey questionnaire requested specific information regarding the purpose of each report form which was identified in Part I. For report forms which were utilized at least once a month, respondents were requested to provide the following information: (1) purpose of the report, including why it is used, who uses it, and for what reason, and (2) when to use the report, including scheduled usage and the circumstances under which the report is completed. In addition to providing the above stated information, respondents were instructed to attach documentary information illustrating or explaining report writing procedures in their departments. These materials included copies of completed forms and reports taken from files which were appended to the questionnaire.

Part III of the survey questionnaire requested that respondents supply ratings for factors in evaluating completed reports. These factors were derived from

an analysis of the completed report forms (see Appendix H) as well as previous related job analysis studies of police patrol officers. Using the five point scale shown below, respondents were asked to evaluate each report writing factor to determine its degree of importance in evaluating the quality of completed narrative field reports.

Rating	Importance
1	Little Importance
2	Some Importance
3	Important
4	Very Important
5	Critically Important

The five factors which were identified included: Clarity, Grammar, Spelling, Detail, and Word Usage. These factors will be discussed in greater detail in the following section of this chapter. The purpose of this aspect of the questionnaire survey was to determine the appropriate relative emphasis to be accorded each writing skills area in the preparation of an examination.

The completed report writing survey forms which constitute the basis for the examination development are included as Appendix I.

Fifty-two (52) law enforcement agencies were contacted for the purpose of soliciting their participation in completing the previously-described survey questionnaire. Of that number, 39 agencies (i.e., 73.6%) responded by returning a questionnaire. Thirty-two (32) questionnaires were properly completed in all respects. The remaining questionnaires were incomplete in one or more aspects. Appendix J is the Response Control Roster prepared by the MLEOTC staff documenting the adequacy of the questionnaires submitted by department. Table 7, which follows, contains the names of the 35 law enforcement agencies upon which the survey data analysis was predicated.

Table 7.

Departments Providing Data Base for Report Writing Survey

1. Adrian Police Department
2. Ann Arbor Police Department
3. Buchanan Police Department
4. Cheboygan County Sheriff's Department
5. Clay Township Police Department
6. Clinton Township Police Department
7. Conrail Police Department
8. Ferris State College Department of Public Safety
9. Flint Police Department
10. Genesee County Parks & Recreation Commission
11. Gladwin County Sheriff's Department
12. Grand Trunk Railroad Police
13. Grosse Ile Police Department
14. Grosse Pointe Shores Department of Public Safety
15. Ingham County Sheriff's Department
16. Kalamazoo Township Police Department
17. Kent County Sheriff's Department
18. Lansing Community College Department of Public Safety
19. Lapeer County Sheriff's Department
20. Lenawee County Sheriff's Department
21. Livonia Police Department
22. Ludington Police Department
23. Michigan State Police
24. Midland Police Department
25. Norton Shores Police Department
26. Owosso Police Department
27. Oxford Police Department
28. Redford Township Police Department
29. Royal Oak Police Department
30. Saginaw County Sheriff's Department
31. Sault Ste. Marie Police Department
32. Tecumseh Police Department
33. Washtenaw County Sheriff's Department
34. Westland Police Department
35. Woodhaven Police Department

Table 8 lists the 16 core writing skills tasks identified in the statewide job analysis study performed by Personnel Research Consultants. For each task, the average number of field reports completed by patrol officers is provided. These data have been derived from the previously described report writing survey forms.

Table 8.

Description of Field Report Usage
by Patrol Officer Job Tasks (N = 35)

Report Writing Requirement	# of Report Forms Which Are Required
Document chain of custody for evidence	2.5
Obtain search warrants and/or make proper return	1.9
Record location of physical evidence at scene	1.8
Record confessions in writing	1.6
Prepare felony complaint forms for warrant authorization	2.3
Summarize . . . statements of witnesses and complainants	1.6
Take statements of witnesses	1.3
Prepare criminal case summary sheet for prosecution	1.4
Prepare misdemeanor complaint forms for warrant authorization	1.9
Transcribe field notes for reports	1.7
Complete standard accident report form (UD-10)	1.7
Write narrative reports	2.0
Complete DUIL arrest reports	2.8
Record circumstances regarding traffic citation	1.6
Complete incident reports by checking boxes, etc.	3.2
Make entries in individual patrol log	1.1
	30.4

An inspection of Tabel 8 on the preceding page indicates the patrol officer in a given month completes an average of 30.4 report forms in conjunction with the 16 identified report writing requirements. The range of reports is substantial and includes both narrative as well as brief checklist type formats. It is the opinion of Wollack & Associates that these data support the usage of a pre-employment writing skills examination for screening job applicants.

Table 9 details the relative importance of the various report writing requirements. The five report writing requirements were:

Clarity: Is the description of persons or events unambiguous and understandable?

Grammar: Is the language of the report grammatically correct?

Spelling: Are the words correctly spelled?

Detail: Is the description of persons or events detailed enough to provide a full account of the facts?

Word Usage: Are the words used correctly?

The previously-described five-point rating scale was utilized for the purpose of determining the relative importance of the various writing skill requirements.

Table 9.

Relative Importance of Report Writing Requirements (N = 32)

Report Writing Requirement	Mean Importance Rating	Percent Weight
Detail	4.94	25
Clarity	4.69	24
Word Usage	3.56	18
Grammar	3.34	17
Spelling	3.22	16
		100%

All of the writing skills factors were seen as being substantially important in the evaluation of field reports. The relative weights of the various report writing factors were derived as percentages of the mean importance ratings which were summed. The factors of "Detail" and "Clarity" are regarded as being most important among the report writing skills. "Word Usage, Grammar" and "Spelling" are regarded as being roughly equally important report writing requirements.

Examination Development

Two item pools of 125 items each were prepared for the purpose of measuring the five writing skills identified in the previous section of this report. The 250 writing skills items consisted of two item types or formats for each of the five writing skills factors. Shown below are the two types of formats for each such factor.

FACTOR	FORMAT A	FORMAT B
Spelling	Item format requires identification of correctly spelled word from among three alternative spellings of same word.	Item format requires identification of incorrectly spelled word from among three different words within a given sentence.
Grammar	Response requires identification of correct grammatical use from among two alternative choices.	Response requires identification of correct grammatical use based upon a single sentence which must be classified as grammatically correct or incorrect.
Clarity	Item format requires identification of clearest phrase or sentence from among three alternatives which are independent	Item format is based upon a subsection of sentence followed by three alternatives, rather than three completely different whole sentences.
Detail	Items require identification of most detailed/complete descriptions among three choices.	Items are composed of a single passage, and call for a judgment of whether the passage is complete or incomplete (i.e., Who? What? Where? When?)

The item formats for the two examination forms were distributed in a random order. The 250 item pool measuring writing skills was administered to the same group of students at Ferris State College and Lansing Community College

described in the previous chapter on the Reading Comprehension Examination. Each student completed the writing skills item pool. Based upon this administration, two 50-item writing skills examinations were prepared. As previously described, an item analysis procedure was utilized to assure the selection of examination items with acceptable item characteristics. In addition, the items were selected in such a way as to conform with the relative weights for the five writing skill item types. These relative weights were determined, as previously noted, from the Report Writing Survey Form, Section III. Table 10 compares the percentage weights of the two examination forms with the job analysis for each writing skills factor.

Table 10.

Relative Weights for the Five Writing Skills Items

Factor	Percentage Weight Job Analysis	Percentage Weight Form A	Percentage Weight Form B
Detail	25	24	24
Clarity	24	24	24
Word Usage	18	18	18
Grammar	17	18	18
Spelling	16	16	16
	100%	100	100%

The job analysis showed, for example, that the factor of "Word Usage" received an 18% weighting indicating the relative importance of this factor in the evaluation of field reports. Therefore, 18% of each of the two examination forms were word usage items (i.e., 9 of 50 items). In this manner, the sampling of the items and their relative emphasis conforms with the job analysis and general requirements for content validation. It may be seen from an inspection of Table 10 that the percentage weights for each of the five job analysis factors closely correspond with the relative emphasis for the five item types in the two examinations.

In addition to developing two writing skills examinations which have an appropriate degree of emphasis in each area, it is necessary to show that the psychometric characteristics of the items are acceptable. Table 11 reports the results of the item analysis for the two writing skills examinations.

Table 11.

Item Characteristics of Two Writing Skills Examinations

Item Characteristics	Form A	Form B
Number of Items	50	50
Mean Item Difficulty Index	24	25
Mean Item Discrimination Index	34	35
K-R 20 Reliability	0.87	0.86
Standard Error of Measurement	2.60	2.69

An inspection of Table 11 shows that each of the two examinations contain 50 items. The mean item difficulty indices for Forms A and B are 24 and 25 respectively, figures which are considered to be very acceptable. The mean item discrimination index for Form A is 34 and for Form B 35. These values are considered to be quite acceptable. The internal consistency reliability of the examinations are 0.87 and 0.86 for Forms A and B respectively. These values are regarded as high levels of reliability given an examination of 50 items in length. The standard error of measurement for Forms A and B respectively are 2.60 and 2.69. In the opinion of Wollack & Associates, the item analysis data are highly satisfactory and describe examinations with very good psychometric qualities. We believe that these data strongly support a conclusion of content validity for both examination forms.

As with the reading comprehension examination, the writing skills tests are regarded to be power tests rather than speeded tests. Time limits to be imposed from the normative analysis would be administrative in purpose only.

Alternate Showing Analysis

The rationale, methodology and sample composition for the alternate showing study was previously described in the Reading Comprehension chapter of this report. The very same sample was utilized for the purpose of performing an alternate showing study with the two forms of the writing skills examinations. It should be understood that the reasoning behind the choice of two writing skills forms is identical to the reasoning in support of two reading comprehension examinations. It was thought that differences in examination formats might

produce more favorable results for some item types with respect to the reduction of adverse impact against minorities.

The previously-described normative sample was utilized to determine examination cutoff scores and the relative impact of both test forms upon the minority officers. Table 12 details the comparison of white and non-white officers on the two forms of the writing skills tests.

Table 12.

Comparison of Writing Skills Test Performance by Racial Groups

Test	Statistics	Whites	Non-Whites
		(N = 166)	(N = 63)
FORM A			
	Mean	42.77	39.40
	Standard Deviation	4.18	4.45
	Acceptance Rates	.90	.71
FORM B			
	Mean	39.99	37.68
	Standard Deviation	4.58	4.80
	Acceptance Rates	.89	.79

Table 12, above, shows the means, standard deviations, and acceptance rates on the two examination forms for the white and non-white officers. For examination Form A, a raw score cutoff of 35 was adopted as compared with a raw score cutoff of 39 on Form B. Both of these cutoffs produced an overall acceptance rate on the writing skills portion of the examination roughly equal to 85%. The rationale for adopting this 85% rule of thumb has been described in the preceding chapter.

Of particular note are the comparative acceptance rates of the two racial groups. For Form A, the white acceptance rate is 90% as compared with 71% for the non-white officers. For this examination form, the non-white officers have an acceptance rate which is equal to 79% of the acceptance rate for the white officers. This finding narrowly falls below the 80% rule of thumb established by the Uniform Guidelines for the purpose of determining adverse impact. For Form B of the

writing skills examination, the acceptance rate for the non-white officers is 79% as compared with 89% for the white officers. For the latter form of the examination, the non-white officers have an acceptance rate which is equal to approximately 89% of the acceptance rate for white officers. This finding for Form B establishes that the use of Form B, with a raw score cutoff of 39, produces no adverse impact against minority officers under the 80% rule. It is, therefore, recommended that Form B of the writing skills examination be adopted for use with the recommended cutoff score.

V. USE OF THE BASIC LITERACY EXAMINATION

To facilitate the administrative ease of the basic literacy examination, it is recommended that the reading comprehension and writing skills tests be combined into a single examination booklet consisting of 100 multiple-choice items. The two sections of the examination should be maintained as separate subtests, and two subtest scores should be generated for the respective portions of the examination. The recommended minimum cutoff scores for the two subtests have been reported in previous chapters. A time limit of 90 minutes is recommended for the 100-item examination. This time limit should be sufficient to allow the vast majority of applicants to complete the examination.

The intent of the basic literacy examination is to identify and deselect those individuals with very substandard capability to deal with written communications. This examination has been developed specifically for the purpose of identifying unqualified applicants. It would be an inappropriate use of the basic literacy examination to attempt to rank order job applicants in order of their test scores. The examination is to be utilized only as a pass-fail instrument. In order to obtain a passing score, the applicant must pass both portions of the basic literacy examination. Should an applicant fail either or both subtests, he or she is regarded as having failed the entire examination.

The racial analyses reported in previous chapters indicated that the method of examination use which has been recommended produced no adverse impact against non-white incumbent police officers in the normative sample. It is generally recognized that the use of examinations in a pass-fail manner produces the least adverse impact against protected classes. Moreover, a pass-fail strategy is considered most appropriate for this examination, because the intent of the MLEOTC is to provide a screening device for identifying individuals with inadequate literacy skills. Therefore, in no case would it be regarded as appropriate to utilize the basic literacy examination in a manner other than that which has been recommended.

It is recognized that in some public agencies, civil service ordinances require the rank ordering of job applicants. Test users are encouraged to use the MLEOTC basic literacy examination only for the purpose of initial screening,

and if rank ordering is required, eligibility should be determined by other means. Test users are encouraged to contact the MLEOTC for assistance in dealing with such problems.

VI. BASIC PHYSICAL SKILLS TEST BATTERY
UNIFORM GUIDELINES: CONTENT VALIDITY

The following index describes the citations, the corresponding requirements, and a listing of reference pages indicating the appropriate sections of this validation report which deal explicitly with the content validation requirements of the Uniform Guidelines on Employee Selection Procedures adopted by the Equal Employment Opportunity Commission, the U. S. Civil Service Commission, the U. S. Department of Labor, and the U. S. Department of Justice (August 25, 1978). This listing includes all reporting requirements which have been designated as being essential for documenting the content validity of employment tests.

Citation	Requirement	Reference
15 C (1)	Dates and location(s) of the job analysis should be shown.	See: A Job Analysis of Police Physical Skill Requirements, 1979, pp. 5-10 Appendices A and D
15 C (3)	A description of the method used to analyze the job should be provided.	Ibid, pp. 1-5 Appendices C and D
15 C (3)	The work behavior(s), the associated tasks, and, if the behavior results in a work product, the work products should be completely described.	Ibid, pp. 17-57
15 C (3)	Measures of criticality and/or importance of the work behavior(s) and the method of determining these measures should be provided.	Ibid, pp. 20-22

Citation	Requirement	Reference
15 C (3)	Where the job analysis also identified the knowledges, skills, and abilities used in work behavior(s), an operational definition for each knowledge in terms of a body of learned information and for each skill and ability in terms of observable behaviors and outcomes, and the relationship between each knowledge, skill, or ability and each work behavior, as well as the method used to determine this relationship, should be provided.	Ibid, pp. 22-57
15 C (4)	Selection procedures, including those constructed by or for the user, specific training requirements, composites of selection procedures, and any other procedures supported by content validity, should be completely and explicitly described or attached.	p. 39 pp. 44-45 Appendix K
15 C (4)	If commercially available selection procedures are used, they should be described by title, form and publisher.	Not applicable
15 C (4)	The behaviors measured or sampled by the selection procedure should be explicitly described.	pp. 45-49
15 C (4)	Where the selection procedure purports to measure a knowledge, skill, or ability, evidence that the selection procedure measures and is a representative sample of the knowledge, skill, or ability should be provided.	pp. 45-49
15 C (5)	The evidence demonstrating that the selection procedure is a representative sample of the work behavior(s), or a representative sample of a knowledge, skill or ability as used as part of a work behavior and necessary for that behavior should be provided.	pp. 39-50

Citation	Requirement	Reference
15 C (5)	The user should identify the work behavior(s) which each item or part of the selection procedure is intended to sample or measure.	p. 49
15 C (5)	Where the selection procedure purports to sample a work behavior or to provide a sample of a work product, a comparison should be provided of the manner, setting, and the level of complexity of the selection procedure with those of the work situation.	pp. 45-49
15 C (6)	The alternative selection procedures investigated and available evidence of their impact should be identified.	pp. 44-58
15 C (6)	The scope, method, and findings of the investigation, and the conclusions reached in light of the findings, should be fully described.	pp. 39-61
15 C (7)	The methods considered for use of the selection procedure . . . and available evidence of their impact should be described.	pp. 51-61
15 C (7)	This description should include the rationale for choosing the method for operational use, and the evidence of the validity and utility of the procedure as it is to be used.	pp. 59-61
15 C (7)	The purpose for which the procedure is to be used (e.g., hiring, transfer, promotion) should be described.	p. 1
15 C (7)	If the selection procedure is used with a cutoff score, the user should describe the way in which normal expectations of proficiency within the work force were determined and the way in which the cutoff score was determined.	pp. 51-58

Citation	Requirement	Reference
15 C (7)	In addition, if the selection procedure is to be used for ranking, the user should specify the evidence showing that a higher score on the selection procedure is likely to result in better job performance.	Not applicable
15 C (8)	The name, mailing address, and telephone number of the person who may be contacted for further information about the validity study should be provided.	William C. Nash MLEOTC 7426 N. Canal Rd. Lansing, MI 48913 (517)-322-1946

VII. PHYSICAL SKILLS TEST

In September, 1979, Wollack & Associates published a comprehensive job analysis study: A Job Analysis of Police Physical Skill Requirements. The findings of that research study clearly suggest the need for a pre-employment physical skills examination to screen police officer job applicants. The remainder of this chapter details the development/validation of such an examination.

Based upon a close review of the aforementioned job analysis report, Dr. Merle Foss developed a comprehensive preliminary physical skills test battery which became the subject of this research study. The battery was comprised of four general types of measures: (1) Static Strength, (2) Cardiovascular Endurance, (3) Work Sample Tests, and (4) Calisthenic Tests. In addition to these four general types of measures, the battery also included a broad range of anthropometric and flexibility measures which were to be utilized in a series of research studies not directly related to the instant study. Table 13 lists the items of the preliminary physical skills test battery. Appendix K contains a detailed description of the examination battery events, equipment utilized, and other specifics pertaining to the methods of measurement.

Table 13.

Preliminary Physical Skills Test Battery

I. STATIC STRENGTH	
Handgrip	10" Lift
90° Press	20" Lift
165° Press	90° Lift
90° Curl	135° Back & Leg Lift
135° Leg Flexion	Shoulder Push
135° Sit-up	Back Push
135° Leg Extension	Arm Push
II. CARDIOVASCULAR ENDURANCE	
Step Test	Half-Mile Run
III. WORK SAMPLES	
Slow Obstacle Course	140 lb. Dummy Push
Fast Obstacle Course	165 lb. Dummy Drag
140 lb. Dummy Drag	95 lb. Carry-Lift
IV. CALISTHENICS	
Chin-ups	60 Seconds Pushups
60 Seconds Sit-ups	

Sample Selection

A sampling plan was developed for the purpose of testing incumbent police officers with the preliminary examination battery. The agencies from which the normative sample was selected are identified in Table 14.

Table 14.

Agencies That Participated in Physical Testing

MICHIGAN STATE POLICE

- Michigan State Police

DETROIT POLICE DEPARTMENT

- Detroit Police Department

LARGE CITIES, VILLAGES AND TOWNSHIPS

- Ann Arbor
- Dearborn
- Dearborn Heights
- Flint
- Grand Rapids
- Lansing
- Livonia
- Pontiac
- Royal Oak
- Southfield
- Warren

MEDIUM CITIES, VILLAGES AND TOWNSHIPS

- Adrian
- Battle Creek
- Bloomfield Township
- Clinton Township
- East Lansing
- Holland
- Jackson
- Michigan State University
- Muskegon
- Redford Township
- Roseville
- Waterford Township
- Wayne State University
- Westland
- Wyoming

SMALL CITIES, VILLAGES AND TOWNSHIPS

- Benton Township
- Buena Vista Township
- Canton Township
- DeWitt Township
- Kalamazoo Township
- Kentwood
- Lansing Township
- Mason
- Melvindale
- Meridian Township
- Owosso
- Woodhaven

LARGE SHERIFFS' DEPARTMENTS

- Berrien
- Calhoun
- Eaton
- Genesee
- Ingham
- Kalamazoo
- Kent
- Livingston
- Muskegon
- Oakland
- Ottawa
- Saginaw
- Washtenaw

Table 14 (contd.)

SMALL SHERIFFS' DEPARTMENTS

- Branch
- Gratiot
- Isabella
- Shiawassee
- VanBuren

RAILROADS

- Grand Trunk & Western
- Norfolk & Western

AIRPORTS

- Capital City

DEPARTMENT OF NATURAL RESOURCES

- Department of Natural Resources

LOCAL PARKS & RECREATION

- Huron/Clinton Metro Authority

Designated departments were telephoned, requesting their participation in the test administration, and the purpose of the testing was explained. A letter was sent to each department that agreed to participate (see Appendix L). When the reply form with the department's choice of officer(s) indicated was received, a letter was sent to each officer (see Appendix M). This letter indicated the purpose of the test, the workshop location, date, and equipment to be brought. Table 15, on the following page, identifies the police officer physical skills test sample. It should be noted that the sample was selected in such a way as to assure representativeness with respect to three criteria: agency type, sex of officer, and police officer age.

Table 15.

Police Officer Physical Skills Test Sample

Agencies	Age Groups 21-25	26-30	31-35	36-40	41+	Totals
1. Michigan State Police						
Male Officers	4	4	5	2	1	16
Female Officers	3	7	5	0	0	15
Totals	7	11	10	2	1	31
2. Detroit Police Department						
Male Officers	2	7	5	2	1	17
Female Officers	1	5	8	3	0	17
Totals	3	12	13	5	1	34
3. Large Cities/Villages/Townships (100+ Officers)						
Male Officers	3	4	4	3	2	16
Female Officers	5	9	2	1	1	18
Totals	8	13	6	4	3	34
4. Medium Cities/Villages/Townships (30-99 Officers)						
Male Officers	1	4	3	3	0	11
Female Officers	10	9	2	0	0	21
Totals	11	13	5	3	0	32
5. Small Cities/Villages/Townships (1-29 Officers)						
Male Officers	2	2	6	0	1	11
Female Officers	3	5	0	0	1	9
Totals	5	7	6	0	2	20
6. Large Sheriff Departments						
Male Officers	3	5	6	1	1	16
Female Officers	3	7	1	1	1	13
Totals	6	12	7	2	2	29
7. Small Sheriff Departments						
Male Officers	0	3	2	0	0	5
Female Officers	0	0	0	0	0	0
Totals	0	3	2	0	0	5
8. Non-Traditional Agencies (Airports, Dept. of Natural Resources, Local Parks and Railroads)						
Male Officers	0	3	1	0	1	5
Female Officers	0	2	1	0	0	3
Totals	0	5	2	0	1	8
GRAND TOTALS						
Male Officers	15	32	32	11	7	97
Female Officers	25	44	19	5	3	96
Totals.....	40	76	51	16	10	193

Normative Testing

A total of 193 officers were tested, of which 97 were males and 96 females. Officers were selected from five age groupings in a manner roughly proportional to their representation in the police patrol officer position. For the purpose of conducting a study of the sexual impact of the examination events, approximately equal samples of males and females were selected. It should be noted that all of the individuals in the normative sample were incumbent officers working the patrol assignment.

In an effort to identify any problems with administration of the testing, a pilot test was conducted on December 8, 1980. Five male officers from the Michigan State Police Training Division, as well as two female civilians from the MLEOTC staff, participated. Test proctors included the Training Council staff and contractual staff supplied by Dr. Foss. Minor problems were identified with regard to the flow of test takers through the testing process, and appropriate revisions were made.

The actual normative testing was conducted at the Michigan State Police Training Academy in accordance with the following schedule:

Date	Agency Type
December 15-18, 1980	MSP and Large Cities, Villages & Townships
January 13-15, 1981	Medium and Small Cities, Villages & Townships
January 19-21, 1981	Large and Small Sheriff Departments
February 3-5, 1981	Detroit Police Department (at Detroit)
February 17-19, 1981	All

All test participants were oriented to the testing through an oral presentation which described the purpose and content of the test battery. Each participant completed a background medical questionnaire, read and signed an Informed Consent Statement, and filled in a Scoring Sheet (see Appendices N, O, and P).

Subsequent to the orientation, subjects were instructed to dress in gym clothing and shoes of their choosing and were processed through the same sequence of seven stations. Appendix Q diagrams the layout of the test stations and the subject flow pattern. A detailed description of the test administration instructions

are contained within Appendix K.

It should be noted that substantial precautions were taken to assure safety. The MLEOTC test coordinator approved subjects to proceed for testing after reviewing the screening test results and considering the comments of the screening technician. Subjects were also provided written instructions about the importance of warming up and stretching prior to being tested; were instructed to cool down after completing the test sequence, and were required to indicate whether they incurred any apparent injuries while being tested. The seven testing stations were manned by a combination of MLEOTC staff, as well as trained measurement technicians from the University of Michigan working under Dr. Foss' supervision.

Selection of Final Test Battery

A test battery consisting of six events was selected from among the 25 physical skills tests administered. This battery was selected on the basis of the following considerations: (1) comprehensiveness in covering the full content domain of police officer physical skills requirements; (2) comparison of male-female performance, and (3) a correlational analysis. The six events comprising the MLEOTC physical skills test battery are:

Combined Handgrip
Slow Obstacle Course
165 lb. Dummy Drag
95 lb. Carry-Lift
60 Seconds Pushups
Half-Mile Shuttle Run

The following is a more detailed description of the examination battery:

Combined Handgrip. Grip strength is measured with the Smedley Dynamometer which measures right and left grip strength in kg units. The values for right and left grip strength are combined.

Obstacle Run. Subject runs forward a distance of 20 feet, drops down and crawls through an obstacle which is 6 feet long and 2.5 feet wide, runs forward 20 feet, climbs over a 6.5 foot barrier using hand/footholds, runs forward 20 feet, runs once around a 2 foot by 4 foot table obstacle following a path which requires him/her to strike an indicator flag at the front top edge of the table twice, runs 20 feet back to the wall and scales it without

the advantage of foot or handholds, drops down from the wall and runs 5 feet to the "stop" pad.

165 Dummy Drag. Subject is timed with respect to ability to drag a 165 lb. life-form dummy a distance of 30 feet.

95 lb. Carry-Lift. Subject is timed regarding ability to carry a 95 lb. duffle bag a distance of 30 feet and to lift and place it upon a table which is 36 inches high.

60 Seconds Pushups. Maximum number of pushups is determined based upon a 60-second time period.

Half-Mile Shuttle Run. Time score is recorded to run a distance of one-half mile which consists of 15 round trips around markers set 88 feet apart.

Evidence of Validity

Three of the examination events are work samples based upon the previously described job analysis findings. Table 16 presents the three work sample events and compares them to the job analysis findings for related types of physical activities. It may be seen from an inspection of Table 16 that a close correspondence exists between the job analysis findings and the examination events.

Table 16.

Comparison of Work Sample Tests to Job Analysis Findings

Job Analysis		Exam Events
LIFTING/CARRYING		95 POUND CARRY-LIFT
Height of Lift:	2.95 feet	Examinee carries a 95 lb. duffle
Distance of Carry:	30.86 feet	bag a distance of 30 feet, and
Weight of Object:	95.37 pounds	lifts and places dummy on a
Officer Unassisted:	69%	table which is three feet high.
Speed Required:	8%	
DRAGGING/PULLING		165 POUND DUMMY DRAG
Distance Moved:	25.67 feet	Examinee drags a 165 lb. human
Weight of Object:	166.92 pounds	form dummy a distance of 30 feet.
Officer Unassisted:	28%	
Speed Required:	34%	
RUNNING/CRAWLING/CLIMBING		OBSTACLE COURSE
Distance of Run:	132 yards	Examinee runs a distance of 105
Obstacles Encountered:	4.20	feet in which he/she crawls through
Height of Barriers Climbed		2.5 foot wide culvert of six foot
Handholds:	6.68 feet	length, twice climbs over a 6.5
Footholds:	7.06 feet	foot wall with hand/footholds,
Speed Required:	26%	and goes around a 2'x4' barrier.
Distance Crawled:	6.78 feet	
Size of Crawlspace:	2-3 feet	
Speed Required:	39%	

The three previously-described events: the 95 lb. Carry-Lift, 165 lb. Dummy Drag, and the Obstacle Course correspond quite well to the job analysis findings with respect to the previously-described quantitative parameters. The remaining three examination events are somewhat more abstract than the work samples, and depend upon further evidence of construct validity. Table 17 presents correlational data for the six-event battery. Of particular note is the correlational evidence for the Static Strength, Cardiovascular Endurance, and Calisthenic measures.

Table 17.

Correlational Data: Six-Event Battery*

Subtests	Combined Grip	Slow Obstacle Course	165 lb. Dummy Drag	95 lb. Carry-Lift	60 Seconds Pushups	Half-Mile Run
I. Static Strength						
Handgrip	--	--	51	66	--	--
90° Press	73	59	62	68	69	--
165° Press	71	58	58	66	69	--
90° Curl	70	55	58	65	53	--
135° Leg Flexion	65	--	57	63	53	--
135° Sit-up	70	53	60	64	54	--
135° Leg Extension	64	--	57	63	53	--
10" Lift	70	--	58	62	55	--
20" Lift	65	--	55	57	55	--
90° Lift	56	--	--	--	--	--
135° Back & Leg Lift	55	--	--	53	--	--
Shoulder Push	61	--	--	55	--	--
Back Push	61	--	--	50	--	--
Arm Push	66	--	56	60	--	--
II. Cardiovascular Endurance						
Step Test	--	--	--	--	--	58
Half-Mile Run	--	50	--	--	50	--
III. Work Samples						
Slow Obstacle Course	--	--	54	57	57	50
Fast Obstacle Course	--	58	70	--	--	--
140 lb. Dummy Drag	55	56	93	69	51	--
140 lb. Dummy Push	--	--	75	66	--	--
165 lb. Dummy Drag	51	54	--	70	51	--
95 lb. Carry-Lift	65	57	70	--	--	--
IV. Calisthenics						
Chin-ups	--	--	--	--	71	--
60 Seconds Sit-ups	--	--	--	--	55	--
60 Seconds Pushups	--	57	51	--	--	50

*Correlations reported only if r equals/exceeds .50

The above table presents correlations only for those measures included in the six-event battery. Moreover, correlations are reported only if the coefficient equals or exceeds r = .50.

The combined handgrip measure is highly correlated with a variety of static strength measures. Obviously, strength is a significant element in a variety of job-related physical activities including: dealing with resisting subjects, pushing and pulling objects, lifting and carrying, etc. The combined handgrip measure appears to be an ideal strength index, in that, it is easily administered with a minimum potential for error. More importantly, this single test shows so high a correlation with the broad range of strength measures that it provides a very useful and valid index of one's static strength. The correlational evidence strongly supports the inclusion of this measure in the examination battery.

The job analysis study of Wollack & Associates revealed many physically-demanding job duties requiring substantial endurance on the part of the officer. Such activities involved running, wrestling with subjects, moving heavy objects, and so forth. It would be appropriate for this reason to include a measure of cardio-vascular endurance in the examination battery. The half-mile shuttle run is such a measure. The data reported in Table 17 indicate a strong correlation between performance on the half-mile run and the step test. The step test was included in the preliminary test battery as an alternate method of measuring cardiovascular endurance. Given the substantial correlation between the two measures, it is recommended that the half-mile shuttle run be included as an event in the examination battery. This suggested event is easily administered, in that, there is no potential for the type of measurement error one might encounter utilizing the step test. Moreover, the half-mile shuttle run is a straightforward, readily administered event which requires no special equipment, and it may be easily monitored.

Sixty (60) seconds pushups were also shown to be a highly valid index of upper body strength and endurance. Exercise physiologists regard this type of measure as being one of "dynamic muscular endurance". The data reported in the previous table tend to bear this out. The 60 seconds pushup measure is substantially correlated with a broad range of static strength measures, the half-mile run, several work samples, and other calisthenic tests. Based on

the strong correlational showing, it is recommended that this event be adopted as part of the final battery, because it provides so valid a measure of upper body strength and muscular endurance.

The work sample measures also proved to be substantially correlated with the static strength measures. In particular, the 165 lb. dummy drag and the 95 lb. carry-lift showed the broadest range of correlations with the static strength measures. The slow obstacle course proved to be substantially correlated with tests covering all four general types of measures. While the validity of the work sample tests does not depend upon a correlational strategy, nevertheless, these intercorrelations do substantially support the inclusion of these types of events in the battery.

Recommended Computational Procedure

The relative emphasis accorded the various elements of a test is an important aspect of content validity. The test user has the burden of showing that the examination content is weighted in a manner which is consistent with the importance and/or criticality of related job duties. The computational procedure for determining one's test score should, therefore, take this factor into consideration. Table 18, which follows, presents a rationale for weighting the examination events.

Table 18.
Weighting of Examination Events

Job Activity	Critical Frequency	Percent Weight	Exam Events
Use of Force	562	22	Grip/Pushups
Lifting/Carrying	400	16	95 lb. Carry-Lift
Running	154	06	Half-Mile Run
Running/Crawling/Climbing	1052	42	Obstacle Course
Dragging/Pulling	351	14	165 lb. Dummy Drag
	2519	100%	

Various types of job activities have been identified in the preceding table based upon the job analysis. Also shown is the "critical frequency" which is merely the total number of related physical incidents, as determined by the job analysis, which were of a critical nature. For example, there were 562 incidents involving the use of force which were regarded as being critical. Using a percentage weight rationale, it is appropriate to assign a 22% weight to those examination events which are intended to measure those physical skills related to the use of force. This would include the Combined Grip Strength measure as well as the 60 Seconds Pushups. Both of these measures, in the aggregate, should receive a 22% weight. Therefore, it is appropriate to weight each of the two measures by 11% in order to achieve an aggregate weighting of 22%. The activity of lifting and carrying objects should receive a 16% weighting, in that, 400 such critical incidents were identified. For the activity of running, 615 critical incidents were identified. Of the activities involving running, approximately one-quarter of those activities involved distances of 100 yards or more. Therefore, the percentage weight for the Half-Mile Run is based upon 25% of 615 incidents or a critical frequency of 154 incidents involving running of more than 100 yards. The relative percentage weight for the Half-Mile Run is, therefore, 6%. The remaining frequency of critical running activities has been combined with the frequencies for crawling and climbing to obtain the total of 1,052 critical incidents involving these activities. It is appropriate to combine these job activities and their associated frequencies, because it is precisely these three physical skills which are being measured in the Obstacle Course. Based upon the critical frequency of 1,052 for the combined job activities, a 42% weight for the Obstacle Course is derived. Finally, a 14% weighting for the 165 lb. Dummy Drag is based upon a critical frequency of 351 for the dragging/pulling of persons.

In computing a composite score for the six-event examination battery, the following procedure was utilized.

1. Each of the six scores was transformed to a T-score with a mean of 50 and a standard deviation of 10.
2. The percentage weights were applied to each of the events in accordance with the weighting scheme described above.
3. A final composite score was calculated by summing the weighted T-scores to form a final score which, itself, is a T-score.

More details will be provided in the following chapter regarding scoring of this examination.

Comparison of Incumbent Officers by Sex

Utilizing this methodology for determining a composite score on the six-event battery, overall scores were determined for the officers in the normative sample. Table 19 presents these data separately for male and female officers.

Table 19.

Comparison of Incumbent Officers by Sex on Physical Skills Test Battery

Composite Scores	Males	%	Cum. %	Females	%	Cum. %
59 - 60	13	14	14	0	0	0
57 - 58	24	25	39	0	0	0
55 - 56	37	39	78	2	2	2
53 - 54	9	9	87	7	8	10
51 - 52	5	5	92	9	10	20
49 - 50	2	2	94	3	3	23
47 - 48	2	2	96	12	13	36
45 - 46	2	2	98	12	13	49
43 - 44	2	2	100	11	12	61
41 - 42	0	0	100	7	8	69
39 - 40	0	0	100	13	14	83
37 - 38	0	0	100	6	7	90
35 - 36	0	0	100	3	3	93
33 - 34	0	0	100	1	1	94
31 - 32	0	0	100	1	1	95
29 - 30	0	0	100	4	5	100
	96	100%		91	100%	

It should be noted, once again, that by definition the T-score has a mean of 50. Bearing this in mind, it is evident from an inspection of the above table that in the combined male-female police officer sample, the vast majority of the male officers score above the mean, whereas, the great majority of female officers

score below the combined mean. The above table describes the distribution of test scores, by sex, giving the number, percentage and cumulative percentage of officers with composite scores falling within the given intervals.

The following table summarizes the preceding table and provides a clearer picture of the magnitude of the disparity between the sexes.

Table 20.

Comparison of Acceptance Rates by Sex for Incumbent Police Officer Sample

Cutoffs	Males	Females
	%	%
57	39	0
55	78	2
53	87	10
51	92	20
49	94	23
47	96	36
45	98	49
43	100	61

This table compares acceptance rates, by sex, for the male and female officers. The term "acceptance rate" pertains to a percentage of each sex group who "pass" the examination given a set of hypothetical cutoff scores. For example, if a cutoff score on the composite were set at 53, it would result in having 87% of the male officers pass the examination as compared with only 10% of the female officers. By looking at a series of hypothetical cutoffs, it is possible to determine for the incumbent police officers what the associated passing rates are. The disparity between the sexes is most evident by considering the higher hypothetical cutoff scores. In particular, if a composite cutoff of 57 were utilized, it would result in a 39% male passing rate as compared with a 0% female passing rate. Even more striking are the results which would pertain given a cutoff of 55. The latter hypothetical cutoff would result in a 78% passing rate for males as compared with a 2% passing rate for females. At the other extreme, a cutoff of 43 would result in a 100% acceptance rate of the male officers as compared with a 61% rate for the females.

To understand better the basis for the disparity in acceptance rates, one must consider the descriptive data by examination event. Table 21 presents such descriptive data for the six-event battery.

Table 21.

Descriptive Data: Six-Event Battery (Incumbents)*

Event	Males	Females	Overall
I. COMBINED GRIP			
Mean (kgs)	120.4	72.5	97.1
s.d.	22.3	14.1	30.5
II. OBSTACLE COURSE			
Mean (seconds)	22.1	48.9	35.3
s.d.	15.7	22.4	23.5
III. 165 lb. DRAG			
Mean (seconds)	5.5	10.3	7.9
s.d.	2.2	4.8	4.4
IV. 95 lb. CARRY			
Mean (seconds)	5.9	46.0	25.6
s.d.	9.4	26.5	28.2
V. 60 SECONDS PUSHUPS			
Mean (repetitions)	25.0	11.4	18.2
s.d.	13.0	8.6	13.0
VI. HALF-MILE RUN			
Mean (seconds)	334.7	402.6	368.1
s.d.	99.9	99.2	105.0

*Officers who were unable to complete an event were assigned a score equal to the most deviant, unfavorable score.

It should be noted in interpreting the data in this table that a statistical procedure was utilized to take account of individuals who were unable to complete an event. Rather than discarding such data, a statistical procedure was applied which, in essence, penalized those officers who were unable to complete the event. The rule of thumb which was utilized was to assign a

score equal to the most deviant, unfavorable obtained score in the event that an officer was unable to complete an element of the examination.

The difference in mean performance between the sexes is notable for each of the six examination events and without exception. In particular, the disparities in mean performance for the obstacle course, 165 lb. dummy drag, and the 95 lb. carry-lift were most conspicuous.

Obviously, the disparities which have been noted are of sufficient magnitude to generate concern on the part of those individuals associated with this test development project. Subsequent to the incumbent normative testing, a substantial effort was made to identify those particular elements of the examination battery which might have accounted for the result with an eye towards a possible redesign or modification of test equipment. As a result of an intensive review of the test content, two events were modified with the hope that female performance might be improved. In particular, the grips for the 95 lb. duffle bag were improved, and the height of the table upon which the bag is to be placed was reduced from 36 inches to 32 inches. It was thought that the latter modification was fully justified, because the 32 inch height tends to be more consistent with the job analysis findings. The second modification involved the 6.5 foot barrier utilized in the obstacle course. In the normative tryout, hand/footholds were used on one side of the barrier, however, the examinee had no such holds on the backside of the wall. It should be recalled that the obstacle course requires the examinee to scale the wall in both directions. Because the solid wall proved to be so substantial a barrier to women, a decision was made to modify the barrier to place hand/footholds on both sides. Moreover, this change also conforms more closely to the job analysis findings.

Administration of the Modified Battery

After these modifications were made, it became necessary to implement the revised examination to determine if the changes had any impact upon the performance of women. A sample of college students at Ferris State College were available for the purpose of this examination tryout. This sample consisted of 29 males and 26 females. It should be noted that the students were in the Criminal Justice Program and had no special background or training

which would render them unrepresentative of the typical job applicant. In particular, it should be noted that these students were not physical education majors.

The modified six-event examination battery was administered by the MLEOTC staff at Ferris State College on April 30-May 1, 1981. The following table provides descriptive data on the revised examination battery, showing results by sex.

Table 22.

Descriptive Data: Revised Test Battery (Students)*

Event	Males	Females	Overall
I. COMBINED GRIP			
Mean	120.9	67.7	95.8
s.d.	14.5	9.5	29.5
II. OBSTACLE COURSE			
Mean	15.4	34.2	24.3
s.d.	3.2	13.2	13.3
III. 165 lb. DRAG			
Mean	5.3	8.6	6.8
s.d.	6.1	2.1	2.2
IV. 95 lb. CARRY			
Mean	4.1	12.2	8.0
s.d.	0.6	8.2	6.9
V. 60 SECONDS PUSHUPS			
Mean	42.9	8.0	26.4
s.d.	18.9	7.3	22.8
VI. HALF-MILE RUN			
Mean	234.5	303.3	267.0
s.d.	20.5	30.4	43.0

*Students who were unable to complete an event were assigned a score equal to the most deviant, unfavorable score.

A number of things are evident from the data in the revised test battery. First, the vast discrepancy which was evident in the test scores of male and female officers on the 95 lb. carry-lift was greatly reduced as a result of the modifications which have been made. On the obstacle course, which had also been modified, test scores for male and female officers alike improved. However, a much greater degree of improvement in the female test scores on the obstacle course was evident. By and large, the most conspicuous differences between the student and incumbent samples were not confined to male-female comparisons. In general, the performance level of the students was considerably superior to that of the incumbent officers. Scores on the two events which had been modified showed considerable improvement as one would expect. Therefore, it is reasonable to attribute such improvement to the modifications rather than to fundamental differences in physical skills between the two samples. However, some clear differences were evident in other elements of the battery which had not been modified. In particular, male students had a mean of 42.9 pushups as compared with a mean of 25.6 for male officers. The difference in performance for the female samples was smaller, but tended to favor the incumbents. On the half-mile shuttle run, the disparity in favor of the student sample was substantial. Male and female students alike had mean scores which were approximately 100 seconds faster than comparable mean scores for the incumbents.

An interesting comparison between the police officer and student samples concerns the percentage of examinees who were unable to complete one or more examination events, i.e., examinees who were unable to scale the 6.5 foot wall in the obstacle course; examinees unable to lift the 95 lb. bag or drag the 165 lb. dummy; examinees who could not do even one pushup, and examinees unable to complete the half-mile run. Table 23 compares the officer and student samples by sex with respect to the percentage of examinees who failed to complete these examination events.

Table 23.
Percent of Examinees Who Failed to Complete Events

Event	Sex	Officers	Students
		%	%
Obstacle Course*	Male	5	0
	Female	44	27
165 lb. Dummy Drag	Male	0	0
	Female	5	0
95 lb. Carry-Lift*	Male	1	0
	Female	50	4
60 Seconds Pushups	Male	2	0
	Females	11	19
Half-Mile Run	Male	7	0
	Female	17	0
Overall	Male	15	0
	Female	73	42

*Event modified for administration to student sample

The male students in the Ferris sample were all capable of completing all of the events. This compares with the incumbent officer sample of males in which a low percentage of such officers were unable to complete the events. In the aggregate, 15% of the male officers were unable to complete one or more of the examination events. The percentage of such female officers was 73% as compared with a 42% figure for the female students. The 73% figure is probably somewhat overestimated, because of the examination modifications which were made following the normative testing. For the 95 lb. carry-lift, the percentage of female officers unable to complete that event was 50% as compared with 4% of the female students. This disparity must largely be attributed to the modification in the design of the event. The same is true for the obstacle course where female performance vastly improved as a result of the modifications. Taking the two modifications into account, it is reasonable to assume that the percentage of examinees in the two female samples who failed to complete the events would be equal. In the view of Wollack & Associates, the most significant conclusion as a result of the two test administrations is that the examination battery, even with the modifications that were made, substantially favors males.

While the elements of the examination battery, we believe, are highly job-related and valid, the impact of this examination upon females is certainly a matter of concern. While a definite effort was made to select subtests which have as little sexual impact as possible, it soon became evident in the test tryouts that any combination of events selected would be apt to produce a comparable degree of adverse impact against women.

VIII. USE OF THE PHYSICAL SKILLS EXAMINATION

The six-event battery described in the preceding chapter of this report is to be used as a screening device for establishing minimal physical standards. As such, one should consider the most appropriate, defensible methods of examination use while bearing this objective in mind. Without a doubt, the method by which an examination is implemented and interpreted is as much a consideration in determining its adequacy and fairness as is the content of the examination.

A Pass-Fail Examination

The MLEOTC physical skills test battery is a basic competency examination and, as such, is to be scored as a pass-fail examination. It is widely recognized that the use of test score data in this fashion generates the least adverse impact against protected classes. This factor is certainly a consideration in determining a strategy for implementing these basic competency examinations. More directly, however, one must consider the primary purpose for this examination. The objective of this test is to identify those candidates who have good physical skills and who are most suited for performing the job duties of a police officer. Moreover, the examination program seeks to eliminate those job applicants whose physical skills are below average and are, therefore, unqualified to perform the job duties of the patrol officer. These considerations justify the adoption of a minimal cutting score below which job applicants are regarded as unacceptable for selection as police officers.

It was previously noted in the preceding chapter that the examination battery was modified slightly as a result of problems identified in the initial testing process with incumbent personnel. The fact that some examination events were changed subsequent to the incumbent testing, strongly suggests the need for additional data gathering to provide a normative basis for determining an examination cutoff. It is recommended that this normative analysis be performed with job applicants rather than incumbent personnel. Certainly, the matter of female impact is of primary consideration in deciding upon an implementation strategy. The use of job applicants as a normative sample would permit a comparison of acceptance rates for actual applicants rather than incumbent personnel.

Wollack & Associates recommends that the examination be implemented for the purpose of data collection to establish cutoffs. On an interim basis, it would be appropriate to utilize the best available information for determining test cutoff scores. It should be recognized that as the data base accrues and the sample becomes larger, it will be possible to establish cutoff scores with a greater degree of certainty than would prevail during the interim period. It is, moreover, assumed that there would be changes in the recommended cutoff scores as additional information is collected.

Method of Test Scoring

The previously reported procedures for scoring the six event battery should prevail during the data-collection phase for actual job applicants. Specifically, each of the six scales should be converted to a standardized score and multiplied by the appropriate weight as determined by the job analysis data. It is recommended that a policy be adopted wherein those applicants who are unable to complete a particular event are assigned, for that event, a test score equivalent to the most negative obtained score. The importance of this procedure is to assure that no individual will be eliminated from the examination on the basis of a failure to complete a single event. This provides an important safeguard, in that, it bases the pass-fail decision upon a composite score rather than a single event. This type of procedure is intended to promote female employment in law enforcement. While it may be justified to eliminate a candidate entirely on the basis of a failure to perform an important examination event, nevertheless, this procedure is not being recommended. In this sense, the examination extends to such individuals who are unable to complete one or more events the opportunity to compensate for such deficiencies by demonstrating stronger performance on the remaining examination measures.

Wollack & Associates strongly believes that the six-event battery is substantially job-related and provides a comprehensive measure of relevant physical skills. Throughout the test development/validation effort, all parties concerned were extremely mindful of the potential adverse impact of any such measures upon female applicants. We regard the representation of females in law enforcement to be a worthwhile and important objective. The examination should not be an absolute barrier to such employment opportunities for women. On the other

hand, a substantial discrepancy in physical performance between the two sexes is evident. This finding is not particularly surprising in view of the self-evident differences in the physical structure of males and females. The preceding job analysis, which provides the basis for this examination, has shown a substantial physical skills requirement in the patrol officer's job. It is, therefore, necessary to counterbalance the objectives of public safety and nondiscrimination in employment. Wollack & Associates believes that the implementation of this examination should be consistent with both objectives.

END