MICHIGAN TEST FOR TEACHER CERTIFICATION

Technical Report

OCTOBER 2014–SEPTEMBER 2015

MICHIGAN Department of Education

OFFICE OF PROFESSIONAL PREPARATION SERVICES (OPPS)
608 WEST ALLEGAN ST, LANSING, MI 48909 | p: 517-373-3310 | f: 517-373-0542

WWW.MICHIGAN.GOV

STATE OF MICHIGAN
STATE BOARD OF EDUCATION

JOHN C. AUSTIN – PRESIDENT
CASANDRA E. ULBRICH – VICE PRESIDENT
MICHELLE FECTEAU – SECRETARY
PAMELA PUGH – TREASURER
LUPE RAMOS-MONTIGNY – NASBE DELEGATE
KATHLEEN N. STRAUS
EILEEN WEISER
RICHARD ZEILE

EX-OFFICIO
RICHARD D. SNYDER – GOVERNOR
BRIAN J. WHISTON – SUPERINTENDENT OF PUBLIC INSTRUCTION
MARILYN SCHNEIDER - EXECUTIVE

MICHIGAN DEPARTMENT OF EDUCATION
608 WEST ALLEGAN STREET • P.O. BOX 30008 • LANSING, MICHIGAN 48909

Office of Professional Preparation Services (OPPS)
www.michigan.gov/teachercert
(517) 373-3310

MICHIGAN STATE BOARD OF EDUCATION
STATEMENT OF ASSURANCE OF COMPLIANCE WITH FEDERAL LAW

The Michigan State Board of Education complies with all Federal laws and regulations prohibiting discrimination and with all requirements and regulations of the U.S. Department of Education. It is the policy of the Michigan State Board of Education that no person on the basis of race, color, religion, national origin or ancestry, age, sex, marital status, or handicap shall be discriminated against, excluded from participation in, denied the benefits of, or otherwise be subjected to discrimination in any program or activity for which it is responsible or for which it receives financial assistance from the U.S. Department of Education.

Copyright © 2016, held by the State Administrative Board, State of Michigan.

Preface

This Technical Report describes test development, updating, administration, scoring, and reporting activities for the Michigan Test for Teacher Certification (MTTC). The activities described in this report began in 1992 and are ongoing. Procedural and technical information is presented for these activities.

The Michigan Test for Teacher Certification program is a collaborative effort accomplished through a close working relationship between the state (represented by the Michigan Department of Education [MDE]) and the contractor (Evaluation Systems group of Pearson). The MDE supervised, reviewed, and approved the plans, activities, and products of the Michigan Test for Teacher Certification that Evaluation Systems group of Pearson provided.

[PAGE INTENTIONALLY LEFT BLANK]

Michigan Test for Teacher Certification

Technical Report

Table of Contents

Preface iii
Executive Summary 1
Chapter 1: Introduction 9
Chapter 2: Test Design 15
Chapter 3: Test Frameworks and Objectives Preparation and Content Validation 23
Chapter 4: Test Item Preparation, Review, and Field Testing 33
Chapter 5: Standard Setting 41
Chapter 6: Test Administration 51
Chapter 7: Test Scoring and Reporting 65
Chapter 8: Annual Program Review 79
Appendix: Sample Score Reports 81
References 89
Cumulative Content Validation Survey Results 91

[PAGE INTENTIONALLY LEFT BLANK]

Executive Summary

Overview

This Michigan Test for Teacher Certification (MTTC) Technical Report provides information on test development, updating, administration, scoring, and reporting activities that have been undertaken since 1992 by the Michigan Department of Education (MDE) and Evaluation Systems group of Pearson (formerly National Evaluation Systems (NES®)). The report highlights the steps taken to establish the validity and reliability of the MTTC program.

The MTTC program is designed to ensure that individuals who wish to become teachers in Michigan schools have the level of knowledge and skills required to perform effectively the job of a qualified Michigan educator in their content area (or student teaching assignment in the case of the Professional Readiness Examination). The test development, updating, administration, scoring and reporting processes employed have been effective in establishing that the MTTC tests are appropriate instruments for that purpose.

The information summarized in this report is designed to show the following.

- The MTTC tests provide important information directly relevant to the licensure of Michigan educators who meet the knowledge and skill requirements mandated by the state.
- The tests are valid—the information they provide is an accurate measure of the knowledge and skills—and have been validated throughout the development of the program.
- The tests are reliable—the information they provide has sufficient consistency to assure the Michigan public that licensure decisions are being made on reasonable grounds.

The Michigan Department of Education and Evaluation Systems have created a testing program designed to meet the unique educational needs of the State of Michigan. Both organizations are committed to refining the program further as it progresses and as the Michigan education environment changes. The people of Michigan can be confident that the testing program will continue to serve them and their children over the years.

Background

Section 1531 of Public Act 451 (1976), as amended by Public Act 267 (1986), Public Act 282 (1992), and Public Act 289 (1995), mandates a testing program as part of Michigan's teacher certification requirements. In June 1991, the Michigan Department of Education awarded National Evaluation Systems a contract to develop and administer the testing program. The purpose of the tests is to ensure that each certified teacher has the level of knowledge and skills required to perform effectively the job of a qualified Michigan educator in their content area (or student teaching assignment in the case of the Professional Readiness Examination). . The tests are not the only basis on which prospective teachers are judged in Michigan, nor are the knowledge and skills covered by the tests the only types of knowledge and skills, or the only professional and personal qualifications, those teachers must have. The tests represent one prerequisite for obtaining a teaching certificate or endorsement in Michigan.

Initial development of the MTTC took place in 1991–1992. During initial development tests were prepared in 76 fields. A Basic Skills (reading, mathematics, and writing) test was developed that is required of candidates seeking a Michigan provisional teaching certificate. Academic content-area tests were developed in 75 fields. The academic content area tests are required of candidates seeking a secondary-level teaching certificate or those teachers in grades 6–8 who teach in specific subject areas. As of Fall 2013, the MTTC program included the Professional Readiness Examination (replacing the Basic Skills test) and 62 tests in various content areas.

A multi-step process involving the participation of Michigan teachers and teacher educators was used to develop the test materials. Pertinent state regulations and policies, curriculum materials, and information from Michigan teacher preparation programs were reviewed to assist in the preparation of test frameworks and objectives. Before the frameworks were finalized, content validation surveys of test objectives for each test field were conducted with additional teachers and teacher educators. Test items were prepared and field tests were conducted with hundreds of students in Michigan. Test items were reviewed by committees of Michigan teachers and teacher educators who participated in Content Validation and Standard Setting activities. The recommendations of Michigan teachers and teacher educators as well as examinee performance on the first operational test form for each test were presented to the State Board of Education. In June 1992, the State Board of Education set the initial passing standards for the MTTC.

Advisory Panels

The MDE constituted two committees, the Teacher Examination Advisory Committee (TEAC) and the Standing Technical Advisory Council (STAC). The TEAC meets on an as-needed basis to advise MDE staff on general policy issues regarding the assessment program. In 2014, the STAC was reconstituted as the Technical Advisory Committee (TAC), which continues to meet annually, or more often if needed, to review MTTC results and to provide advice on technical issues.

Ongoing Test Development and Updating Activities

An important feature of the MTTC program is an ongoing review and updating of test materials. These periodic reviews are designed to ensure that the testing program continues to address the knowledge and skills needed by Michigan teachers and to reflect current practice in Michigan schools. Test updating began in 1992. Each year the Michigan Department of Education and Evaluation Systems determine the fields that will be reviewed and updated. The updating process has included the following steps:

- Collection and review of Michigan policy documents and curriculum materials;
- Development of test objectives that appropriately reflect Michigan curriculum and classroom practice;
- Review of draft test frameworks (including subareas, objectives, and descriptive statements) by committees of Michigan teachers and teacher educators;

- Conduct of Content Validation Surveys of Michigan teachers and teacher educators to evaluate the importance of the content of the test objectives to the job of a Michigan educator;
- Development of test items that appropriately reflect Michigan curriculum and classroom practice;
- Review of draft test items by committees of Michigan teachers and teacher educators;
- Field testing of new items;
- Standard setting activities with Michigan teachers and teacher educators; and
- Determination of passing standards by the Michigan Department of Education.

Test Validation Process

The "validity" of a test refers to the ability of the test to support the inferences that are to be drawn from it. Test validation is the process of gathering evidence that the test measures what it is supposed to measure and that the inferences to be made from the test scores are supported by evidence. Various types of evidence may be considered in establishing the validity of a test, and a number of methods are typically used to gather such evidence.

The focus of the validation efforts for the MTTC was on establishing that the content measured by the tests was directly related to the specific knowledge and skills required for teacher certification in Michigan. Materials were collected and reviewed from various sources, including approved Michigan teacher preparation programs; state certification statutes, regulations, and minimum teacher qualifications; state curriculum regulations and guidelines; curriculum materials used in Michigan classrooms; textbooks in use in Michigan classrooms and Michigan teacher preparation programs; professional publications; MDE curriculum guides and goals; and Michigan student assessment materials.

Several levels and types of validity evidence were gathered. Validity evidence was gathered pertaining to the overall purpose and goals of the testing program, the actual test content to be measured, the structure and content of the test objectives measured by each test, and the specific test questions measuring each objective. In this way, validation was a central part of each step in the development of the tests, proceeding from the general level of program authorization to the specific level of the test questions themselves.

The validation process for the MTTC followed professionally accepted procedures for the validation of certification tests, including the following steps.

- The purpose of the testing program was established by the mandating legislation.
- Test objectives—which define eligible test content—were based on Michigan laws, Michigan standards for teacher certification, and other state documents/policies provided by the Michigan Department of Education.
- The test objectives were reviewed by MDE staff, a Michigan Bias Review Committee, and a Michigan Content Advisory Committee for each test field. Revisions were made based on the feedback received.
- For each test field, a Content Validation Survey of the proposed test objectives was conducted among Michigan school teachers and college faculty at Michigan colleges and universities with approved teacher preparation programs. The survey asked Michigan educators to rate the importance of the test objectives to the job of a Michigan school teacher. Test objectives rated important were eligible for inclusion on the tests.
- Draft test items for each field were reviewed by a Michigan Bias Review Committee and a Michigan Content Advisory Committee.

- Draft test items were field tested on operational test forms and with students enrolled in Michigan teacher preparation programs.
- Michigan school teachers and college faculty participated in item validation and standard setting activities. Test items were reviewed to ensure that they matched the validated test objectives and were accurate, free from bias, and related to the job of an educator in Michigan.

In these ways, the content of the tests—from the most general level to the most specific—was repeatedly and systematically judged by qualified reviewers to be valid for the MTTC program.

Bias Prevention

Prevention of bias in the MTTC is important as a matter of fairness and as an aspect of validity. Guarding against bias in the MTTC materials involved the collaboration of educators and reviewers focused on excluding language, content, or perspectives that might disadvantage examinees based on background characteristics irrelevant to the purpose of the test, and on including content and perspectives that reflect the diversity of the Michigan population.

The review of MTTC materials from the standpoint of bias prevention is not a single activity that is conducted at a particular point in the development process. Rather, it is an ongoing aspect of all components of the development process, from the definition of test content through the review of all test items that may appear on all test forms.

Educators from diverse backgrounds were invited to participate throughout the development and updating activities of the MTTC. They served as members of Content Advisory Committees, reviewing draft test frameworks and items, and serving on item validation and standard setting panels. Bias review was a responsibility of every educator who reviewed test materials. In addition, a separate Bias Review Committee, composed of a diverse group of Michigan educators, had the specific responsibility of examining test materials for potential bias.

Standard Setting

Committees of Michigan educators reviewed test questions in their fields of expertise and made judgments concerning the level of knowledge required to perform the job of an entry-level teacher in Michigan. Their item-level judgments were compiled to determine a recommended passing score for each test. The Michigan Department of Education considered the committee recommendations in setting the passing score for each test.

Reliability of Test Results

Reliability refers to the consistency of scores received by a group of examinees. The Appendix presents a number of statistical estimates of the reliability for scores on the tests in the program. For the MTTC tests, as with other tests used in a licensing environment, a relevant reliability statistic is one that provides an estimate of the decision consistency of the tests (i.e., the consistency of the pass/fail decisions that examinees receive).

The meaning of specific reliability estimates derived from the scores of a given group of examinees on a given occasion is a matter of interpretation and judgment that depends on several factors, including the nature of the tests for which the estimates are provided and the purpose for which the results will be used.

Statistical estimates of reliability may be affected by several factors related to the test itself and the group of examinees on which they are based. Key factors that influence reliability estimates include the following:

- *Number of examinees*—In general, reliability estimates based on larger numbers of examinees are more stable than estimates based on smaller numbers.
- *Test length*—Reliability estimates tend to be higher for tests with greater numbers of questions.

- *Composite tests*—Reliability estimates for tests with both multiple-choice and performance components (i.e., Professional Readiness Examination (formerly the Basic Skills test) and Spanish) are more meaningful when they are based on the combined, total test than on either of the two components alone.
- *Test content*—Reliability estimates are typically higher for tests that cover narrow, homogeneous content than for tests, such as teacher certification tests, that cover a broad range of content.
- *Examinees' knowledge*—Reliability estimates tend to be higher if examinees in the group have widely varying levels of knowledge and lower if they tend to have similar levels of knowledge.

Administration

The MTTC tests are administered under standardized, consistent procedures at sites across Michigan. Test administrations are designed to provide a professional, equitable, and secure testing environment for candidates, including candidates with needs for alternative testing arrangements. Test sites are screened and selected based on criteria relating to test security, accessibility, and appropriate testing conditions and facilities.

Test administrators are oriented and provided with procedural manuals to use before and during the test administration. The orientation process and the manuals are designed to facilitate secure, efficient, and professional test administrations for all candidates.

As of December 2010, computer-based testing became available as an alternative to paper-based testing for some fields. The following table indicates the availability of computer-based testing as of fall 2015. Candidates are able to test at any one of a number of VUE Pearson Professional Centers (PPCs) and authorized test centers in Michigan, as well as at any of hundreds of PPCs throughout the United States, and in over 165 countries. Computer-based testing supplemented the four paper-based test administrations offered during the 2014-2015 program year; candidates had access to two-week testing windows in each month during which a paper-based administration is not offered. During a testing window, testing is available every day except Sunday and select holidays.

Availability of Computer-based Testing

	12/2010	12/2011	12/2012	12/2013	11/2014	05/2015
103 Elementary Education (formerly 083)	x	x	x	Updated		
196 PRE (formerly Basic Skills) - Reading	x	x	x	x		
296 PRE (formerly Basic Skills) - Mathematics	x	x	x	Updated		
396 PRE (formerly Basic Skills) - Writing	x	x	x	Updated		
022 Mathematics (Secondary)		x	x	x		
089 Mathematics (Elementary)		x	x	x		
002 English			x	x		
090 Language Arts (Elementary)			x	x		
106 Early Childhood Education			x	x		
009 History				x		
084 Social Studies (Secondary)				x		
093 Integrated Science (Elementary)				x		
105 Social Studies (Elementary)				x		
044 Physical Education					x	
051 School Counselor					x	
056 Cognitive Impairment					x	
059 Emotional Impairment					x	
063 Learning Disabilities					x	
017 Biology						x
018 Chemistry						x
028 Spanish						x
075 Bilingual Education						x
094 Integrated Science (Secondary)						x

Conclusion

This Technical Report presents information concerning test development, updating, administration, scoring, and reporting activities conducted for the Michigan Test for Teacher Certification program since 1992. The report reviews the steps that were taken by the Michigan Department of Education and Evaluation Systems to establish and verify the validity and reliability of the tests included in the MTTC program.

The Remainder of This Report

The following chapters of this report describe the major steps in test development, updating, administration, scoring, and reporting activities of the Michigan Test for Teacher Certification program from 1992 through September 2014. Individual chapters describe in greater detail the major features of the project.

Chapter 1 provides an introduction to the report, including:

- a summary of key aspects of the development and updating process for the tests; and
- a discussion of the validation process used for the tests.

Chapter 2 describes the test design of the Michigan Test for Teacher Certification, including:

- the sources of the test design;
- the design of the tests included in the MTTC; and
- the structure of the tests in the program.

Chapter 3 describes the preparation and validation of test objectives, including:

- the nature and purpose of the test framework, test objectives, and descriptive statements;
- the basis of the test objectives; the review and validation of test objectives through reviews by the Bias Review Committee and Content Advisory Committees; and
- the conduct and analysis of the Content Validation Surveys of Michigan school teachers and teacher preparation faculty at Michigan institutions of higher education.

Chapter 4 describes the preparation, review, bias review, and field testing of test items, including:

- the preparation of items in the various formats required by each test design;
- the review of items by the Bias Review Committee;
- the review of items by the Content Advisory Committees; and
- field testing of proposed items and item formats.

Chapter 5 describes the procedures used to validate and set standards for the tests, including:

- the validity verification and standard setting meetings; and
- the analysis of the standard setting meeting results.

Chapter 6 describes test administration policies and procedures, including:

- planning and implementing the registration process, including registration for alternative testing arrangements; and
- implementing test administrations at sites across Michigan.

Chapter 7 describes test scoring and reporting, including:

- scoring the tests; and
- preparing results reports for candidates, the Michigan Department of Education, and Michigan teacher preparation institutions.

Chapter 8 describes the annual program review process.

[PAGE INTENTIONALLY LEFT BLANK]

Chapter 1: Introduction

The Michigan Test for Teacher Certification (MTTC) program was developed and is administered on behalf of and under the supervision of the Michigan Department of Education (MDE). The MTTC program is one component of the overall teacher certification process in Michigan. Section 1531 of Public Act 451 (1976), as amended by Public Act 267 (1986), Public Act 282 (1992), and Public Act 289 (1995), mandates that candidates for teaching certificates in Michigan pass a test of basic skills (reading, writing, and mathematics) and in the subject area(s) in which they seek endorsement(s). The purpose of the MTTC program is to ensure that candidates for teacher certification demonstrate the level of knowledge and skills required to perform effectively the job of a qualified Michigan educator in their content area (or student teaching assignment in the case of the Professional Readiness Examination).

The MTTC tests are used in the process of educator certification. Certification tests have particular characteristics that distinguish them from other tests; they are developed and validated according to guidelines and procedures appropriate to the licensure and certification context. The MTTC program was developed under those guidelines.

Initial development of the MTTC took place from the summer of 1991 through the summer of 1992. During that time period, Evaluation Systems (formerly known as National Evaluation Systems) worked with the Michigan Department of Education (MDE) to develop tests in 76 fields. A Basic Skills (reading, writing, and mathematics) test, required of candidates seeking a Michigan provisional teaching certificate, was developed. Content-area tests were developed in 75 content fields. Candidates must pass an academic content-area test in the field(s) in which they seek endorsement(s).

An important feature of the MTTC program is an ongoing review and updating of test materials. In some cases, the review has called for the development of new test fields. In other cases, existing test fields are reviewed and updated to ensure that the content continues to reflect the knowledge and skills needed by Michigan teachers. Additionally, fields may be retired in keeping with updated certification regulations. As of fall 2013, the MTTC program included the newly implemented (in fall 2013) Professional Readiness Examination (formerly the Basic Skills test). There are currently 62 tests in various content areas.

In this introduction, a summary of the key aspects of the development and updating process for the tests is described. The schedule for test development and updating is presented. Finally, the validation process used in the development and updating activities is presented.

Test Development and Updating Process

Periodic reviews of the MTTC are designed to ensure that the testing program continues to address the knowledge and skills needed by Michigan teachers and to reflect current practice in Michigan schools. Test review and updating began in 1992 following the completion of initial development activities. Each year the MDE determines the fields that will be reviewed and updated. The test development and updating process typically takes place over a two-year period.

In the first year, Evaluation Systems works with the MDE to collect and review Michigan policy documents and curriculum materials for the relevant test fields. This step is undertaken to ensure that the tests continue to be grounded in Michigan law, policy, and practice. Following a review of the materials, Evaluation Systems develops a set of draft test objectives for each test field that appropriately reflects Michigan curriculum and classroom practice. The test objectives are formatted into a test framework for each field. Committees of Michigan teachers and teacher educators review the appropriate test framework(s) to ensure that they are accurate and appropriate for the MTTC program. They are also reviewed to ensure that they are free from potential bias and represent the diversity of the Michigan population.

After review of the test frameworks by committees of Michigan teachers and teacher educators, each test framework is formatted into a Content Validation Survey. Content Validation Surveys are distributed to appropriate samples of Michigan school teachers and college faculty who prepare teachers. Survey respondents rate the importance of the content of the test objectives to the job of a Michigan teacher.

During the first year of the test development and updating process, test items reflecting Michigan curriculum and classroom practice are prepared to match each objective in the test framework. Michigan school teachers and teacher educators review the draft test items to ensure that they are accurate and appropriate for the MTTC. As with the test framework, draft test items are also reviewed to help ensure that they are free from potential bias and represent the diversity of the Michigan population.

Field testing is completed during the second year of the test development and updating process. Draft test items that were reviewed and approved by Michigan educators are eligible for field testing. Draft items may be field tested as nonscorable (i.e., not included in the calculation of examinee scores) on appropriate MTTC operational test forms. Draft items may also be field tested at separate field tests conducted with students enrolled in Michigan colleges and universities with approved teacher preparation programs.

At the end of the second year of the test development and updating process, committees of Michigan teachers and teacher educators meet again to provide judgments about the performance on the test items on the first new test form of individuals who have the level of skills and content knowledge required to perform effectively the job of a qualified Michigan educator in their content area (or to perform effectively in their student teaching assignment, in the case of the Professional Readiness Examination). These standard setting judgments are provided to the MDE, which then sets the passing standard for the test. Evaluation Systems implements the MDE passing standard for each test field, completing the test development and updating process.

Schedule for Test Updating Activities

The table on the following pages presents the list of test fields that have been updated or redeveloped since the initial round of MTTC test development. Fields marked with an asterisk (*) indicate new tests that were added to the original set of MTTC fields.

Validation Process

Test validity is defined as the extent to which a test measures what it is designed to measure. In the case of the MTTC, validity refers to the extent to which the tests measure the basic skills (i.e., reading, writing, and mathematics) and academic content area knowledge and skills needed to perform the job of an educator in Michigan schools. Throughout the test development and updating activities of the MTTC, careful attention was paid to ensuring the validity of the tests.

The validation process for the MTTC followed professionally accepted procedures for the validation of certification tests. The standards for educational and psychological testing, published by the American Psychological Association (1985 and 1999), Equal Employment Opportunity Commission (EEOC) regulations, and judicial decisions (such as the 1996 decision in a case brought in California against the California Basic Educational Skills Test™ [CBEST®], a teacher certification test of reading, writing, and mathematics skills) suggest that the appropriate validation strategy for a licensure and certification test is the use of content-based validation.

Development Periods **(test available as of Fall of final year of development period)**	**Test Fields** Fields marked with an asterisk (*) indicate new tests that were added to the original set of MTTC fields.	
1992–1994	Basic Skills Elementary Education*/** Industrial Arts Language Arts	Mathematics Reading* Science
1993–1995	Biology History	Psychology Social Studies*
1994–1996	Early Childhood Education English	Health Spanish
1995–1997	Business Education Chemistry Guidance Counselor	Learning Disabled Physical Education
1996–1998	Emotionally Impaired Geology/Earth Science	Speech
1997–1999	Library Media Mentally Impaired	Middle Level* Physics
1998–2000	English as a Second Language (ESL)* Industrial Technology*	Social Studies Technology and Design*
1999–2001	English Family and Consumer Sciences Journalism	Mathematics (Secondary) Mathematics (Elementary)* Physical Education
2000–2002	Communication Arts (Secondary)* Computer Science	Language Arts (Elementary)* Reading Reading Specialist*
2001–2003	Geography Health History	Visual Arts Education (formerly Art Education)
2002–2004	Political Science Biology	Integrated Science (Secondary)* Integrated Science (Elementary)*

Development Periods **(test available as of Fall of final year of development period)**	Test Fields (continued)	
2003–2005	Chemistry Earth / Space Science (formerly Geology / Earth Science)**	Economics Physics Basic Skills: Writing (review)
2004–2006	Business Management, Marketing, and Technology* (formerly Business Education)	Marketing Education Music Dance Physical Science*
2005–2007	Spanish French Japanese*	Learning Disabilities (formerly Learning Disabled) Cognitive Impairment (formerly Mentally Impaired)
2006-2008	Autism Spectrum Disorder (formerly Autism) English as a Second Language	German Chinese (Mandarin)* Emotional Impairment (formerly Emotionally Impaired)
2007-2009	Physical or Other Health Impairment Bilingual Education	Latin Russian Arabic (Modern Standard)*
2008-2010	Speech	
2008-2012	Early Childhood Education (General and Special Education)	
2008-2013	Basic Skills (Mathematics) (renamed the Professional Readiness Examination: Mathematics subtest in fall 2013) Elementary Education	
2009 – 2013	History Social Studies (Elementary)* Social Studies (Secondary)	
2010-2013	Economics Geography Political Science	
2011-2013	Basic Skills (Writing) (renamed the Professional Readiness Examination: Writing subtest in fall 2013) School Counselor (formerly Guidance Counselor)	
2015 (Re-set Standards only)_	Autism Spectrum Disorder Cognitive Impairment Emotional Impairment Learning Disabilities	

*Fields indicating new tests that were added to the original set of MTTC fields.
**Tests for which passing standards were reset in 2012 with the updated description of the hypothetical candidate.

A content-based approach to validation is appropriate for the MTTC because the tests are designed to measure specific skills and knowledge (i.e., those skills and knowledge required for teacher certification in Michigan as specified in state legislation). The content required of each test is specified in the test objectives and test items. The focus of validation efforts for the MTTC involves ensuring that the test objectives and test items are consistent with Michigan statutes, regulations, educational practice, and reflect the knowledge and skills judged important for the job of a Michigan teacher.

Development of test objectives. A set of test objectives that describes the test content was developed for each test field. The test objectives were derived from Michigan state standards and Michigan program, policy, and curriculum materials. These Michigan sources serve as a link between the tests and Michigan requirements. The test objectives were reviewed by committees of Michigan educators and validated as important for Michigan teachers through Content Validation Surveys.

Content validation surveys. For each test field, a content validation survey was conducted to verify that the test objectives were important to the job of a Michigan teacher. School teachers and college faculty who prepare teachers participated in the surveys. Results from the surveys were analyzed and the test objectives in each field were found to be important to the job of a Michigan teacher.

Content review and validation of test items. For each field, draft test items were prepared and reviewed by committees of Michigan teachers and teacher educators. A Bias Review Committee of Michigan educators reviewed draft items to ensure that they were free from bias and representative of the Michigan population. Content Advisory Committees reviewed items to ensure that they matched the validated test objectives, and were appropriate, accurate, free from bias, and related to the job of a Michigan teacher.

Field testing. After draft test items were reviewed and approved by Michigan educators, they were field tested to gather information on test item performance. Field testing occurred during operational administrations of the MTTC and at Michigan colleges and universities with approved teacher preparation programs. Test item performance was reviewed at the conclusion of the field tests.

Standard setting. Committees of Michigan educators met again to provide judgments of the performance on the test items of individuals who have the level of skills and content knowledge required to perform effectively the job of a qualified Michigan educator in their content area (or to perform effectively in their student teaching assignment, in the case of the Professional Readiness Examination). These judgments were used by the MDE to set the passing standard for each test.

Passing standards. The Michigan Department of Education used the professional judgments of the committees of Michigan educators in determining the passing standard for each test.

[PAGE INTENTIONALLY LEFT BLANK]

Chapter 2: Test Design

This chapter provides information on the test design for the MTTC tests. The MDE and Evaluation Systems worked together to design the MTTC during initial test development activities in 1991. A design was selected for each test that would reflect Michigan laws and regulations governing teacher certification and school teaching. Test development and updating activities have been guided by the initial decisions regarding test design.

The tests included in the MTTC are designed to measure knowledge of a body of content (i.e., professional readiness/basic skills, content areas) defined through a test framework. Each test framework is composed of a set of test objectives. Individually, each test objective describes one aspect of the content of the test field. Together, the set of test objectives is called a test framework and defines the certification field for testing purposes. Test objectives are elaborated by descriptive statements and are grouped into related areas of content called subareas.

Test frameworks are organized into subareas, objectives, and descriptive statements for two reasons. First, this organization is useful in guiding test development. Second, the organization of the objectives into subareas is useful for reporting examinee scores. Examinees, teacher preparation institutions, and the MDE receive information on examinee performance in each subarea. This information is helpful to examinees and faculty in identifying content in which further study or preparation may be beneficial.

Test Design

Originally, the Michigan Test for Teacher Certification program comprised 76 tests: one Basic Skills test and 75 academic content area tests. As the program has developed, some tests have been added and some have been deleted to ensure that the program continues to meet legislative and policy requirements within Michigan. As of fall 2013, the MTTC consisted of the Professional Readiness Examination (formerly the Basic Skills test) and content-area tests in 62 test fields. This design reflects Michigan state regulations governing teacher certification.

Evaluation Systems prepared initial draft test designs after reviewing various state and federal regulations and guidelines. Among these were:

- curriculum guides and essential goal statements for various subject areas (Essential Goals and Objectives, Standards of Quality);
- MDE policy and administrative rules for special education; and
- materials related to the Michigan Educational Assessment Program (MEAP).

Evaluation Systems also reviewed materials obtained from Michigan teacher preparation programs to assist in the preparation of the tests.

Design of the Professional Readiness Examination (formerly the Basic Skills Test). In fall 2013 the Professional Readiness Examination (PRE) was introduced. As an updated version of the Basic Skills test, the PRE includes three subtests: reading, mathematics, and writing. Reading and mathematics are assessed through multiple-choice questions. The Writing subtest in the PRE was redesigned. In the Basic Skills test, writing skills were assessed through a writing assignment in which the examinee prepares an organized, developed composition in edited English in response to instructions regarding content, purpose, and audience. In the PRE, writing skills are assessed both through multiple-choice items in combination with two items requiring candidates to prepare written responses.

The reading subtest includes 42 multiple-choice test questions, of which 36 are scorable. The mathematics subtest includes 45 multiple-choice test questions, of which 40 are scorable. The writing subtest consists of 42 multiple-choice test questions, of which 36 are scorable, and two constructed-response questions. To pass the Professional Readiness Examination, an examinee

must pass each of the subtests (or meet alternative passing measures as determined by the MDE). Once an examinee passes a subtest, she or he does not have to take that subtest again. The Technical Advisory Committee endorsed the updated design of the Professional Readiness Examination.

Redevelopment activities for the Basic Skills (which was renamed the Professional Readiness Examination in 2013) mathematics subtest began in 2008 and the writing subtest began in 2011. The mathematics subtest, when implemented in 2013, retained the original structure as described above. The writing subtest, as of fall 2013, includes 42 multiple-choice test questions, of which 36 are scorable, as well as two writing assignments. The new subtests became operational in October 2013, at which time the test was renamed as the Professional Readiness Examination.

Design of the academic content-area tests. The academic content-area tests are designed to measure content area knowledge and skills. The following table indicates the number of items, both scorable and nonscorable, on each test as of fall 2014.

Code	Field Name	MCQs (All)	MCQs (Scorable)	CRIs
2	English	100	80	--
3	Journalism	100	80	--
4	Speech	100	80	--
5	Reading	100	80	--
7	Economics	100	80	--
8	Geography	100	80	--
9	History	100	80	--
10	Political Science	100	80	--
11	Psychology	100	80	--
12	Sociology	100	80	--
17	Biology	100	80	--
18	Chemistry	100	80	--
19	Physics	100	80	--
20	Earth/Space Science	100	80	--
22	Mathematics (Secondary)	80	64	--
23	French	80	64	2
24	German	80	64	2
26	Latin	80	64	2
27	Russian	35	28	8
28	Spanish	80	64	2
29	Italian	100	80	--
36	Marketing Education	100	80	--
37	Agricultural Education	80	64	--
40	Family and Consumer Sciences	100	80	--
43	Health	100	80	--
44	Physical Education	100	80	--
46	Dance	100	80	--
48	Library Media	100	80	--
50	Computer Science	100	80	--

Code	Field Name	MCQs (All)	MCQs (Scorable)	CRIs
51	Guidance Counselor	100	80	--
53	Fine Arts	100	80	--
56	Cognitive Impairment	100	80	--
57	Speech and Language Impaired	100	80	--
58	Physical or Other Health Impairment	100	80	--
59	Emotional Impairment	100	80	--
61	Visually Impaired	100	80	--
62	Hearing Impaired	100	80	--
63	Learning Disabilities	100	80	--
64	Autism Spectrum Disorder	100	80	--
75	Bilingual Education	100	80	--
83	Elementary Education	100	80	--
84	Social Studies (Secondary)	100	80	--
85	Middle Level	100	80	--
86	English as a Second Language	100	80	--
87	Industrial Technology	100	80	--
88	Technology and Design	100	80	--
89	Mathematics (Elementary)	80	64	--
90	Language Arts (Elementary)	100	80	--
91	Communication Arts (Secondary)	100	80	--
92	Reading Specialist	100	80	--
93	Integrated Science (Elementary)	100	80	--
94	Integrated Science (Secondary)	100	80	--
95	Visual Arts Education	100	80	--
97	Physical Science	100	80	--
98	Business, Management, Marketing, and Technology	100	80	--
99	Music Education	100	80	--
100	Japanese	35	28	8
101	Chinese (Mandarin)	35	28	8
102	Arabic (Modern Standard)	35	28	8
103	Elementary Education	150	80	--
105	Social Studies (Elementary)	100	80	--
106	Early Childhood Education (General and Special Education)	100	80	--
196	Professional Readiness Exam: Reading Subtest	42	36	--
296	Professional Readiness Exam: Mathematics Subtest	45	40	--
396	Professional Readiness Exam: Writing Subtest	42	36	2

Test item formats. Multiple-choice questions and constructed-response items are used in the MTTC program. For the multiple-choice test questions, there are four response alternatives per question, one of which is the best answer of the choices given. Examinees are asked to select the one best answer to each test question and record their response on a scannable answer sheet. There is no penalty for guessing.

Constructed-response items are included on the writing subtest of the Professional Readiness Examination and on the World Language tests and the Latin test. The constructed-response items on the Professional Readiness Examination require an examinee to prepare an organized, developed composition in edited English in response to instructions regarding content, purpose, and audience that are included in the test question. Responses are evaluated on the basis of the following criteria: appropriateness, focus and unity, organization, development, and grammar and conventions. Examinees record their response to the constructed-response item(s) on an answer document (in PBT) or in the space provided for the response on the computer.

The Spanish, French, German, and Latin content-area tests each contain two written constructed-response items. The written constructed-response items each require an extended written response in the target language to a written stimulus that is presented in the test booklet. Examinees record their responses to the written constructed-response items on answer documents that are provided with the test materials.

The Chinese (Mandarin), Arabic (Modern Standard), Russian, and Japanese tests each contain eight constructed-response items, including:

- two listening comprehension assignments that each require the examinee to listen to recorded excerpts and then prepare a written response, in either English or the target language, to the assignments presented in the test booklet;
- two reading comprehension assignments that each require the examinee to read a passage and prepare a written response, in either English or the target language, to the assignment that follows;
- two writing assignments that each require an extended written response in the target language to a written stimulus that is presented in the test booklet;
- a language structures section that requires the examinee to demonstrate a command of the structures and usage of the target language; and
- a language comparisons section that requires a written response, in either English or the target language, comparing and contrasting a given element of English with that element in the target language.

Structure of the Content-Area Tests

While the content covered by each content-area test included in the MTTC program is different, the structure of the content of the tests is similar. In general, the structure is as follows:

- Each content-area test field is organized into about three to six subareas that define the major content-area knowledge and skills of the test.
- Each subarea includes three or more test objectives. The test objectives are broad, conceptual statements, written in language that reflects the skills, knowledge, and understanding that a teacher needs in order to teach in Michigan schools.
- Each test objective is clarified and further described by descriptive statements, which provide examples of the types of knowledge and skills covered by the test objective.
- Each multiple-choice question is written to assess understanding of a particular test objective.
- The number of multiple-choice items that address a given subarea on each test is generally proportional to the number of test objectives within that subarea. For example, a subarea with 25 percent of the objectives is measured by approximately 25 percent of the scorable items on the test form. The approximate percentage of items on each subarea of each test is listed in the study guide for each test field.
- Any one test form comprises a sampling of items measuring the content of the field.
- Constructed-response items are written to assess understanding of a test objective.

The chart that follows illustrates the relationship among subareas, test objectives, descriptive statements, and test items in a typical test field; it is followed by a sample page from one test framework that shows each element of a test framework (i.e., subarea, test objective, and descriptive statement).

Organization of Content within a Sample Test Field

Example	**Description**
Elementary Education	**Test Field** This is the name and code number of the test field.
English Language Arts and World Languages (Subarea I)	**Subarea** Each test field is divided into major content areas within the test field that reflect an organizational principle that is understandable to individuals studying and working in the field. The number of test objectives within each subarea may vary, depending on the breadth of content contained within the subarea.
Understand the major concepts, principles, and instructional practices in the acquisition and learning of languages to create opportunities for communication in a multilingual global society. (Test Objective 001)	**Test Objective** Each subarea contains several test objectives that define content knowledge Michigan educators determined to be important to the job of an educator in this field. Test objectives are broad, conceptual statements that reflect some of the skills, knowledge, and understanding needed by educators in Michigan schools.
Includes knowledge of language as a dynamic system, and strategies for helping students acquire and use language in listening, speaking, reading, and writing for social and academic purposes. (Descriptive Statement for Test Objective 001)	**Descriptive Statement** Each test objective is further elaborated by descriptive statements that provide examples of the range of knowledge and skill included within the test objective. The examples that are provided in descriptive statements are not inclusive of all content that may be covered.
Which of the following activities, performed by young children while a story is being read to them, would most likely encourage them to listen for meaning? A. tapping their feet to the rhythm of the words B. miming the action of the story C. raising their hands when they hear the main character's name D. making a clay sculpture of their favorite character	**Multiple-Choice Item** Each multiple-choice item corresponds to one test objective and typically presents candidates with introductory information, a statement or question to be answered, and a choice of four responses, one of which is the best choice of the responses given.* *The correct response to the sample item, which corresponds to the Sample Test Objective, is B.

Field 83: Elementary Education

Test Objectives (Sample)

Language Arts

Mathematics

Social Studies

Science

The Arts

Health and Physical Education

SAMPLE

SAMPLE: SUBAREA I

LANGUAGE ARTS

0001 Understand the development of reading competence, including interactions among reader, text, and context.

Includes the development of emergent literacy in young children; factors affecting readers' construction of meaning through interactions with text (e.g., readers' prior knowledge; nature, genre, structure, and features of text; context of the reading act); and knowledge of different comprehension strategies for different purposes (e.g., reading a text book to review for a test, reading for enjoyment).

0002 Use vocabulary skills (e.g., structural analysis, contextual analysis) to determine meaning in given passages, and apply knowledge of vocabulary skills to reading.

Includes the use of word structure (e.g., phonetic analysis, syntactic cues, affixes) and context clues to determine the meaning of unfamiliar words; the use of context clues to determine the intended meaning of a word with multiple meanings; and recognition of ways in which figurative language (e.g., metaphor) is used in a given text.

* * *

0010 Understand communication through the listening process.

Includes processes of audio perception and discrimination; attending to messages; assigning meaning; evaluating messages, responding to messages; and remembering message content.

[PAGE INTENTIONALLY LEFT BLANK]

Chapter 3: Test Frameworks and Objectives Preparation and Content Validation

This chapter provides information concerning the preparation and validation of test objectives of the MTTC program.

Evaluation Systems prepared test frameworks—sets of test objectives—to fulfill the design of each test. The objectives are based on Michigan standards, curriculum guides, textbooks, and Michigan teacher education and certification standards.

The test objectives were reviewed by a committee of Michigan bias reviewers appointed by the MDE and by content advisory committees of Michigan school teachers and teacher educators appointed by the MDE. Further validation evidence was gathered through the conduct of Content Validation Surveys for each test field by Michigan school and college educators.

Test Objectives

Organization. The tests included in the MTTC are objective-based in that the content measured by each test is defined by a set of test objectives. As described in the previous chapter, the test objectives are organized into about three to six subareas, or major domains of content, that facilitate scoring and reporting. Within each subarea, a set of test objectives defines the content of the test. Each test objective consists of two parts:

1. the *objective statement*, which broadly defines the content that a teacher in Michigan needs to know, and
2. the *descriptive statement*, which provides examples of the types of knowledge and skills covered by the test objective.

Purposes. The purposes of the test objectives are to:

- establish a link between test content and Michigan legal, policy, and regulatory sources;
- communicate to candidates for certification the general content of the tests and suggest areas of knowledge that the candidate's educational preparation should include;
- communicate to college and university faculty responsible for the preparation of teachers the general content of the tests;
- communicate to the general public and legislators how the standards and expectations for teachers in Michigan are embodied in the MTTC program; and
- provide a context for the reporting and interpretation of test results.

Validation Basis of Test Objectives

The test objectives for each field were designed to reflect Michigan educational statutes, regulations, standards, and policies, as well as school educational practice and, where appropriate, relevant national educational standards.

The test objectives were designed to meet the parameters described below.

- The test objectives describe skills and content-area knowledge that are consistent with Michigan educational practices and curricula in the schools.
- The test objectives describe the knowledge and skills that are important to the job of a Michigan teacher.

Validation through Reviews of Test Objectives

Several activities were conducted to gather validity evidence for the test objectives. Michigan school teachers and college and university faculty who prepare teachers reviewed the test objectives to ensure that they were accurate and appropriate for Michigan teachers. Additional reviewers examined the test objectives to help ensure that they were free from potential bias and representative of the diversity of the Michigan population.

Validation review of test objectives by members of the Michigan Bias Review Committee. Validation of certification tests typically entails a review of test materials for potential bias. Freedom from bias is regarded as an aspect of test validity in that the performance of examinees on a test that is free of bias can be considered to reflect their actual knowledge and skills, rather than irrelevant background characteristics or other factors unrelated to the purpose of the test.

Bias Review Committee. A Bias Review Committee (BRC), selected by the MDE, includes Michigan teachers and teacher educators. The BRC is a standing committee that was formed in 1991 during the initial development of the MTTC program. The BRC reviewed the test objectives and their associated descriptive statements to ensure that they were free from bias and reflective of the diversity of the Michigan population.

Initially, to identify potential BRC members, the MDE sent nomination forms to school districts, the teacher union, and institutions of higher education with approved teacher preparation programs in Michigan. Upon receipt of nominations, the MDE approved a list of educators who could provide representation of both school teachers and teacher educators, representation of educators from across the state, and representation of diverse groups of educators.

Since 1991, the MDE has solicited additional nominations for BRC committee members. Evaluation Systems has assisted the MDE in these efforts. For each bias review of test objectives, the MDE and Evaluation Systems invited the BRC members to attend the review session. The invitation provided background information on the MTTC program, described the responsibilities of the BRC, and provided information regarding the logistics of the review conference.

Conference Introduction. At the conference, the MDE provided an overview of the background, purpose, and policies of the MTTC program. Evaluation Systems oriented and trained committee members for the review task, facilitated the conference, and recorded the recommendations of the BRC in the master copies of the review materials.

Orientation and Training. BRC members were given an overview of the purposes of their review activities and a description of the procedures for the review session. The training included a description of the structure and function of the subareas, objectives, and descriptive statements; the procedures and materials to be used in their review of the test frameworks; and the review criteria to be applied. Committee members were invited to raise questions during and after the orientation and training session and throughout the review activities.

Materials and Procedures. BRC members were sent the book *Fairness and Diversity in Tests*, produced by Evaluation Systems, which outlines issues that relate to bias reviews of test materials. They were instructed to bring this book to the meeting and to refer to it, as necessary, during their reviews. Extra copies of this book were made available at the conference and continue to be made available at each conference. At the conference, committee members were also given the test frameworks for the fields that they were to review.

Committee members were asked to review the test objectives based on the following criteria, which were updated during the 2013-14 program year to reflect updated Michigan expectations:

Content: Does any element of the objectives or descriptive statements contain content that disadvantages a person because of her or his gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Language: Does the language used to describe any element of the objectives or descriptive statements disadvantage a person because of her or his gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Offense: Is any element of the objectives or descriptive statements presented in such a way as to offend a person because of her or his gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Stereotypes: Does any element of the objectives or descriptive statements contain language or content that reflects a stereotypical view of a group based on gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Diversity: Does the list of objectives and descriptive statements permit appropriate inclusion of content that reflects the diversity of the Michigan population?

BRC members reviewed each subarea, objective, and descriptive statement with reference to the bias review criteria. They reached consensus regarding the need for revisions. Recommendations for revisions were recorded by an Evaluation Systems facilitator in the master copy of each test framework. The recommendations of the BRC were presented to the appropriate Content Advisory Committee (CAC) for their consideration. Each CAC was instructed to address bias-related comments made by the BRC.

Validation review of test objectives by Michigan Content Advisory Committees. Validation of certification tests entails, in addition to a review for potential bias, the alignment of test content with professional practice in the field. For the MTTC, the alignment of the test content (via the test frameworks and objectives) with Michigan educational practice in the schools was confirmed by reviews of the test materials by Michigan school teachers and college and university faculty who prepare teachers. A Content Advisory Committee for each test field met to examine the test frameworks (including the subareas, objectives, and descriptive statements) to help ensure that the content is accurate and appropriate for testing teacher certification candidates in Michigan.

Content Advisory Committees. Michigan educators reviewed the test objectives and their associated descriptive statements for content accuracy and appropriateness for teacher certification in Michigan. The MDE selected a Content Advisory Committee (CAC) for each test field. Each CAC included Michigan teachers and teacher educators. For test fields that were undergoing updating activities, individuals who previously served on CACs were invited to participate again. The MDE solicited additional committee nominations for these fields, as needed, aided by efforts made by Evaluation Systems. Upon receipt of nominations and applications, the MDE approved a list of educators that included school teachers and teacher educators from various geographical, ethnic, and racial backgrounds with content-specific expertise in the particular field that they were reviewing.

The MDE and Evaluation Systems invited the selected educators to attend a CAC review session. The invitation provided background information on the MTTC program, described the responsibilities of the CAC, and provided information regarding the logistics of the review conference.

Conference Introduction. At the conference, the MDE provided an overview of the background, purpose, and policies of the MTTC program. Evaluation Systems oriented and trained committee members for the review task, provided each CAC with the recommendations from the BRC, facilitated the conference, and recorded the recommendations of the CAC in the master copies of the review materials.

Orientation and Training. CAC members were given an overview of the purposes of their review activities and a description of the procedures for the review session. The training included a description of the structure and function of the subareas, objectives, and descriptive statements; the procedures and materials to be used in their review of the test frameworks; and the review criteria to be applied. Committee members were invited to raise questions during and after the orientation and training session and throughout the review activities.

Materials and Procedures. The members of each committee read the test framework (including subareas, objectives, and descriptive statements) to become familiar with their structure. Committee members were asked to review the test framework and objectives based on the following criteria (with the criterion for "Freedom from Bias" updated in the 2013-14 program year).

Framework Review Criteria

Program Purpose

- Is the framework consistent with the purpose of the MTTC (i.e., to determine whether prospective teachers have the knowledge and skills to perform the job of an educator in Michigan)?

Organization

- Is the framework organized in a reasonable way?
- Are the subarea headings accurate and do they clearly describe the content?

Inclusiveness

- Is the content of the framework complete?
- Does the framework reflect the knowledge and skills an educator should have in order to teach the content?
- Is there any content that should be added?

Objective Review Criteria

Significance

- Do the objectives describe knowledge and skills that are important for educators to have?

Accuracy

- Do the objectives accurately reflect the content as it is understood by educators in the field?
- Are the objectives stated clearly and accurately, using appropriate terminology?

Freedom from Bias

- Are the objectives free from elements that might potentially disadvantage an individual because of her or his gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Job-Relatedness

- Do the objectives cover important knowledge and skills that an educator should have in order to perform the job of a Michigan educator?

CAC members reviewed each subarea, objective, and descriptive statement with reference to the Framework and Objective Review Criteria. They were provided with the recommendations of the BRC for their test field. CAC members reached consensus regarding the need for revisions. Recommendations for revisions were recorded by an Evaluation Systems facilitator in the master copy of each test framework.

Post Conference Revisions. Following the BRC and CAC reviews of the draft test frameworks, Evaluation Systems incorporated committee revisions as approved by the MDE. Revised test frameworks were prepared for the Content Validation Survey.

Content Validation Surveys

In addition to the validity information gathered about the test objectives from the bias and content reviews described above, Evaluation Systems gathered additional validity information through subsequent surveys of Michigan teachers and teacher educators. Content validation serves to establish a link between test content and the requirements of teaching in Michigan (i.e., job relatedness). The results of the content validation surveys were used to validate the set of objectives that were eligible to be measured in each field.

Content validation survey invitations were sent by mail to principals for distribution to teachers following MDE approval of BRC and CAC revisions to the draft test frameworks. Each survey invitation provided a secure web site address and a user name and password. Respondents logged into the secure site to complete the content validation survey. Respondents to the content validation surveys included currently certified and practicing school teachers in Michigan. For the content-area tests, currently certified and practicing Michigan school teachers who held an endorsement in the test field were eligible to be included in the Content Validation Survey sample. Respondents rated the importance of each test objective in the framework to the job of a Michigan teacher in that subject area.

Teacher educators at colleges and universities in Michigan with approved teacher education programs were also surveyed. For the content-area tests, teacher educators were asked to review each test objective and indicate the importance of the objective for teaching in the subject area field in Michigan schools.

Survey sample design for school teachers. Up to 200 practicing teachers in each field were sampled for the surveys of the content-area test objectives. For high incidence fields such as the Professional Readiness Examination (formerly Basic Skills test) and Elementary Education test, up to 400 practicing teachers were sampled. For lower incidence fields, such as Latin, for which there were fewer than 200 practicing teachers in the field, all teachers were included in the sample. School teachers were sampled from data files provided by the MDE. The data files included assignment area(s), certificate(s), and endorsement(s) for each Michigan school teacher.

Evaluation Systems used the data files to determine the eligible population for each test field. The primary population eligible for each test field consisted of individuals holding Michigan teaching endorsements and teaching assignments corresponding to the test field. Certified teachers holding part-time positions were considered job incumbents and eligible to participate in the survey. In cases in which the primary population was smaller than the survey target number, a secondary population was identified as eligible. The secondary population consisted of individuals holding a corresponding teaching endorsement.

A stratified random sampling design was used to select the school teacher samples from the MDE data files of the population of teachers eligible to participate. The sample was stratified on the variable of race/ethnicity in terms of three categories: African American, white, other.

To meet expectations established with the MDE for representation of minority educator survey participation, Evaluation Systems oversampled African Americans for each test field at roughly twice the rate at which they were represented in the populations of teachers for the test field. This guideline applied in fields in which sufficient numbers of teachers were available.

If teachers were eligible members of more than one test field (because of multiple teaching assignments and certification endorsements), they were selected to participate in only one test field. This was done to focus the attention of the respondent in one test area and to increase the likelihood of an individual's responding to the survey. Sampling was done first for the lowest incidence fields to achieve an adequate sample size for the low incidence fields.

Survey sample design for teacher educators. Teacher educators were sampled from colleges and universities offering approved teacher education programs in the appropriate test fields. Using MDE materials, Evaluation Systems identified the Michigan colleges and universities offering approved teacher education programs in each test field. For each institution identified, Evaluation Systems worked with a MTTC campus contact person to distribute the content validation survey materials.

Teacher educators at Michigan institutions with approved teacher education programs were eligible to participate. To be eligible to complete a survey for a content-area test field, eligible faculty members had to be teaching education courses or academic specialization courses in the content field being surveyed.

MTTC campus contacts were sent the survey invitations and selected the teacher educators to participate. The number of survey invitations sent to an institution for a given test field depended upon the program size at each institution. The MTTC campus contact was provided with written instructions outlining the steps required to prepare a sampling list and select a sample for each field.

Survey instrument design. Separate survey instruments were developed for each of the groups (school teachers and teacher educators) included in the survey. Each survey included the following elements.

General survey instructions. At the beginning of the survey, Evaluation Systems provided general instructions for completing the survey. These described the estimated length of time to complete the survey, how to fill out the survey response form, how to make comments, and how and when to return materials. The confidentiality of survey responses was also explained.

Eligibility and background information questions. Respondents were asked a question to determine their eligibility to participate (e.g., for school teachers, whether they were certified to teach in Michigan in the survey field). If respondents answered "No" to the eligibility question, they were asked to stop and return all materials to Evaluation Systems.

Eligible respondents were asked to complete a series of background information questions. Included were questions regarding teaching assignment(s), gender, race/ethnicity, type of certificate, and level of education.

Objective rating instructions. Respondents were asked to answer a question about each objective using their judgment to determine what they considered important for teaching in the designated field in Michigan schools. The following objective rating questions were used.

School teachers: *In your job as a Michigan teacher, how important is the objective to an understanding of the content of this endorsement area?*

Teacher educators:

To a person preparing for a job as a Michigan teacher, how important is the objective to an understanding of the content of this endorsement area?

Respondents were asked to rate each objective using the following five-point scale.

1 = no importance
2 = little importance
3 = moderate importance
4 = great importance
5 = very great importance

Beginning in 2005, two additional questions were posed to both school teachers and teacher educators. The first question pertains to the descriptive statements. The second question pertains to the set of objectives as a whole.

"How well does the set of descriptive statements represent important examples of the knowledge and skills addressed by the objective?"

"How well does the set of objectives, as a whole, represent important aspects of the knowledge and skills required for performing the job of an entry-level Michigan teacher in this endorsement area?"

Participants were asked to respond to each question using the following five-point scale.

1 = poorly
2 = somewhat
3 = adequately
4 = well
5 = very well

Comments/suggestions. This section was provided for respondents to make comments about the survey, the objectives, or to note any objectives that should be added.

Professional Readiness Examination: Writing Subtest [formerly **Basic Skills (Writing)] 2012.** For the content validation survey for the Basic Skills (currently the Professional Readiness Examination) Writing subtest conducted in the Spring of 2012, the survey sample design and survey instrument design were prepared to address the specific target candidate population and content that were the subject of the subtest. For the Basic Skills test, which is required of all teacher candidates, eligible respondents included all currently certified and practicing school teachers in Michigan. In all cases, respondents rated the importance of each test objective in the framework for a candidate beginning his or her student teaching requirement in Michigan.

Survey sample design. For the Basic Skills (Writing) survey, target sample sizes were 400 practicing teachers and 200 practicing teacher educators. Teacher-preparation faculty members who taught undergraduate or graduate courses to OR prepared undergraduate or graduate education candidates were considered eligible to respond. Surveys were distributed such that 75% of the survey booklets were designated for distribution to education faculty, and 25% of the survey booklets were designated for English or Language Arts content faculty.

Survey instrument design. For the Basic Skills (Writing) survey, the general survey instructions and the eligibility and background information questions mirrored those of the surveys for the subject-area tests.

Objective rating instructions. School teacher and teacher educator respondents were asked to answer a question about each objective using their professional judgment to determine what they considered important for candidates beginning their student teaching requirement in Michigan. The following objective rating question was asked.

"*How important is the knowledge or skill represented by this objective for candidates to have before beginning their student teaching requirement in Michigan?"*

Respondents were asked to rate each objective using the following five-point scale.

1 = no importance
2 = little importance
3 = moderate importance
4 = great importance
5 = very great importance

In addition, the following questions were posed to both school teachers and teacher educators. The first question pertains to the descriptive statements. The second question pertains to the set of objectives as a whole.

"*How well does the* ***set*** *of descriptive statements represent important examples of knowledge and skills addressed by the objective?"*

"*How well does the* ***set*** *of objectives, as a whole, represent important examples of writing knowledge and skills required for candidates to have before beginning their student teaching requirement in Michigan?"*

Participants were asked to respond to each question using the following five-point scale.

1 = poorly
2 = somewhat
3 = adequately
4 = well
5 = very well

Comments/suggestions. This section was provided for respondents to make comments about the survey, the objectives, or to note any objectives that should be added.

Survey distribution and collection: practicing teachers. Initially, Evaluation Systems distributed the school teacher surveys to the selected teachers at school addresses. The survey materials included cover letters from the MDE and Evaluation Systems, a survey booklet, a response form, and a postage-paid return envelope.

The sampled teachers were asked to return their completed surveys by a specified deadline using the postage-paid envelope provided. Evaluation Systems monitored survey returns. Teachers who did not return the surveys by the specified deadline were sent a follow-up survey mailing. The follow-up survey mailing included a follow-up cover letter, a second survey booklet, a response form, and a postage-paid return envelope. The cover letter specified a final due date for return of completed surveys.

Beginning in 2012 with the content validation survey for Basic Skills: Writing subtest and School Counselor test fields, while the process to contact sampled educators remained the same, online administration of the survey was implemented.

Survey distribution and collection: teacher educators. An advance notification letter was sent to MTTC campus contacts at colleges and universities offering approved teacher education programs. The letter provided background information about the program and the content validation survey invitations, informed the campus contacts of their responsibility for overseeing the distribution of the survey invitations and was accompanied by a preliminary survey inventory for the institution. Evaluation Systems distributed the teacher educator survey invitations to the campus contacts along with guidelines for distributing the survey invitations and an inventory of survey invitations sent. The survey packages included a cover letter, and a survey invitation. Postage-paid return envelopes were provided for any unused invitations. The cover letter provided background information about the program, described the purpose of the survey, provided the secure site, user name and password for accessing the content validation survey, and emphasized the importance of the respondents' participation in the survey.

The MTTC campus contacts used the guidelines for distributing the survey invitations to select faculty members to participate. The sampled faculty members were asked to complete the online survey by a specified deadline using the postage-paid envelope provided. Evaluation Systems monitored survey returns. Evaluation Systems identified and called any institutions from which few or no surveys had been completed. If needed, duplicate sets of survey invitations were sent to the institutions. Evaluation Systems extended the deadline for the completion of surveys for institutions that received and/or distributed the information late.

Data preparation. As surveys were completed online, they were logged by type of respondent (i.e., school teacher or teacher educator). An "ineligible" survey response was defined as a response on which the respondent filled in the circle corresponding to "No" for the eligibility question or left the eligibility question blank. An "eligible" survey response was defined as a response on which the respondent filled in the circle corresponding to "Yes" for the eligibility question. Final response rates were calculated as the number of "eligible" returns divided by the total number of surveys sent less the number of "not used" and "ineligible" returns.

Data analysis. For each test field, the following reports were prepared for each sample (school teacher and teacher educator).

Demographic Summary Report. For the background information questions included in the survey booklet, Evaluation Systems generated frequency distributions (number and percent) for each question.

Rating Summary. The rating summary provided, for each objective in the test field, the number of eligible respondents; the arithmetic mean importance rating; the standard deviation of the importance rating; the standard error of the mean; the grand mean, standard deviation, and standard error across all objectives; and the distribution of responses on the 1–5 rating scale. The analyses for the Rating Summary (school teachers) were statistically weighted to generate population parameter estimates that took into account possible oversampling by race/ethnicity.

Outcomes. The data from the Content Validation Survey for each field were analyzed and prepared for presentation to the MDE for review. For surveys conducted through 2004, a criterion score of 2.5 was used as the minimum acceptable mean rating for determining the validity of each objective. Beginning in 2005, the criterion was raised to 3.0 at the request of the MDE. Evaluation Systems worked with the MDE to review and identify any objectives or descriptive statements that did not meet the criterion. The decision for retaining or deleting objectives or descriptive statements rested with the MDE. The MDE had the option of deciding to retain objectives or descriptive statements that did not meet the criterion score. For example, the MDE could decide to retain objectives that did not meet the criterion but, nevertheless, included content that should be eligible for testing (e.g., emerging content in the field, content that is related to a job requirement, content that is significant for the purpose of providing diversity, and/or content that is related to health and safety issues).

The School Teacher and Teacher Educator Survey Results table in the Appendix shows the year that each content validation survey was completed, the number of survey respondents for each test field, and the mean importance rating for all test objectives in each test field. The number of survey respondents and the mean importance rating for all test objectives in each test field is reported separately for school teacher and teacher educator survey respondents.

Chapter 4: Test Item Preparation, Review, and Field Testing

This chapter provides information concerning the preparation, review, and field testing of the test items. As with the development of the test objectives, the importance of providing validity evidence was emphasized throughout the test item preparation process.

For each test field, Evaluation Systems prepared items to meet the approved test design and to correspond to the approved test objectives. Following the completion of item preparation, the draft items were reviewed by the Bias Review Committee and the appropriate Content Advisory Committee. The Bias Review Committee considered the materials to help ensure that they were free of potential bias and representative of the diversity of the Michigan population. The Content Advisory Committees reviewed the test materials to help ensure that they were accurate and appropriate for Michigan school teachers.

Field testing was conducted to gather information about the characteristics and performance of the test items. Items were field tested on operational test forms during regularly scheduled MTTC administrations, in separate rooms during operational test administrations, and at scheduled sessions at Michigan educator preparation institutions.

Test Item Preparation Process

The following process was used to prepare test items for review by the Michigan Bias Review Committee and the field-specific Michigan Content Advisory Committee.

- Evaluation Systems reviewed the test design to determine the number and types of items that would be needed for the test field.
- Evaluation Systems reviewed the revised test framework and test objectives to become familiar with the content that would be assessed by the test items.
- Evaluation Systems reviewed existing sets of test items for correspondence to the revised test framework and test objectives or developed new items, as needed, to match the test framework and test objectives. Item development teams considered the following criteria in the review of the items: objective match, accuracy, freedom from bias, and job-relatedness. Evaluation Systems also reviewed the statistical characteristics of the existing test items.
- Items were prepared for review by Michigan educators.

Item Review by Michigan Educators

Following item preparation, test items were reviewed by the Michigan Bias Review Committee as well as the field-specific Michigan Content Advisory Committee.

Review of test items by members of the Michigan Bias Review Committee. As previously noted, validation of tests in the licensure and certification context typically entails a review of test materials for potential bias. The Michigan Bias Review Committee was formed with the specific mission of reviewing test materials to help ensure that they are free from bias and representative of the Michigan population. The BRC review of test items for potential bias and diversity is one step in establishing the validity of the testing program.

Bias Review Committee. Individuals previously selected by the MDE to serve on the MTTC Bias Review Committee as well as new participants who met the criteria were invited by the MDE and Evaluation Systems to participate in the item review conference. The invitation provided background information on the MTTC program, described the responsibilities of the BRC in the review of test items, and provided information regarding the logistics of the review conference.

Conference Introduction. At the conference, the MDE provided an overview of the background, purpose, and policies of the MTTC program. Evaluation Systems oriented and trained committee members for the review task, facilitated the conference, and recorded the recommendations of the BRC in the master copies of the review materials.

Orientation and Training. BRC members were given an overview of the purposes of their review activities and a description of the procedures for the review session. The training included a description of the structure and function of the test items; the procedures and materials to be used in their review of the test items; and the review criteria to be applied. Committee members were invited to raise questions during and after the orientation and training session and throughout the review activities.

Materials and Procedures. BRC members were provided with the test frameworks for the test items they were reviewing. They were also provided with the draft test items and the review criteria to be applied.

Bias Review Committee members were asked to review the draft test items based on the following criteria, updated in the 2013-14 program year:

Content: Does the item contain content that disadvantages a person because of his or her gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Language: Does the item contain language that disadvantages a person because of his or her gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Offense: Is the item presented in such a way as to offend a person because of his or her gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Stereotypes: Does the item contain language or content that reflects a stereotypical view of a group based on gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

Diversity: Taken as a whole, do the items include content that reflects the diversity of the Michigan population?

BRC members independently reviewed each test item with reference to the bias review criteria. They noted any potential bias in the items. Then, the BRC discussed items a group and reached consensus regarding the need for revisions to any test items. Recommendations for revisions were recorded by an Evaluation Systems facilitator in the master copy of the test items. The recommendations of the BRC were presented to the Content Advisory Committee (CAC) for their consideration. Each CAC was instructed to address bias-related comments made by the BRC.

Review of test items by Michigan Content Advisory Committees. Following the review of the draft test items by the Michigan Bias Review Committee, field-specific Michigan Content Advisory Committees met to examine the test items to help ensure that the content is accurate and appropriate for testing teacher certification candidates in Michigan. This review added to the validation evidence for the MTTC program by helping to ensure that the test content is aligned with educational practice in Michigan schools.

Content Advisory Committees. Individuals who had been selected to review the test framework and test objectives for the test field were invited by the MDE and Evaluation Systems to participate in the Item Review Conference. Additional individuals who met the criteria for participation were also identified and, pending MDE approval of qualifications, invited to participate. The invitation provided background information on the MTTC program, described the responsibilities of the CAC, and provided information regarding the logistics of the review conference.

Conference Introduction. At the conference, the MDE provided an overview of the background, purpose, and policies of the MTTC program. Evaluation Systems oriented and trained committee members for the review task, provided each CAC with the recommendations for the BRC, facilitated the conference, and recorded the recommendations of the CAC in the master copies of the review materials.

Orientation and Training. CAC members were given an overview of the purposes of their review activities and a description of the procedures for the review session. The training included a description of the structure and function of the test items; the procedures and materials to be used in their review of the test items; and the review criteria to be applied. Committee members were invited to raise questions during and after the orientation and training session and throughout the review activities.

Materials and Procedures. Committee members were provided with the test framework for their field, the draft test items, and the item review criteria to be applied. Committee members were asked to review the draft test items according to the following criteria, which were updated during the 2013-14 program year to reflect updated Michigan expectations:

<u>Item Review Criteria</u>

Objective Match

- Does the item measure an important aspect of the test objective?
- Is the level of difficulty appropriate for the testing program?
- Are the items, as a whole, consistent with the purpose of the MTTC program?

Accuracy

- Is the content accurate?
- Is the terminology in the item correct and appropriate for Michigan?
- Is the item grammatically correct and clear in meaning?
- Is the correct response accurately identified?
- Are the distractors plausible yet clearly incorrect?
- Are the stem and response alternatives clear in meaning?
- Is the wording of the item stem free of clues that point toward the correct answer?
- Are the graphics (if any) accurate and relevant to the item?

Bias

- Is the item free of language, content, or stereotypes that might potentially disadvantage or offend an individual based on gender, sexual orientation, gender identity or expression, race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?
- Are the items, as a whole, fair to all individuals regardless of gender, sexual orientation, gender identity or expression race, nationality, ethnicity, religion, age, disability, or cultural, economic, or geographic background?

- As a whole, do the items include content that reflects the diversity of the Michigan population?

Job-Relatedness

- Is the content job-related?
- Does the item measure content or skills that an educator needs on the job in Michigan schools?
- Does the item measure content or skills that an educator should be expected to know in order to perform effectively the job of a qualified Michigan educator (i.e., not learned on the job)?

CAC members independently reviewed each item with reference to the Item Review Criteria. They were provided with the recommendations of the BRC for their test field. CAC members reached consensus regarding the need for revisions. Recommendations for revisions were recorded by an Evaluation Systems facilitator in the master copy of the draft test items.

Post Conference Revisions. Following the BRC and CAC reviews of the draft test items, Evaluation Systems incorporated committee revisions as approved by the MDE. Revised test items were prepared for field testing.

Field testing

Purpose. The purpose of field testing is to gather information about new test items and to help verify that they have acceptable statistical and qualitative characteristics; that is, to:

- determine item difficulty and other psychometric characteristics of the test items before their use as scorable items on operational test forms; and
- evaluate the appropriateness of constructed-response items (Professional Readiness Examination (formerly Basic Skills test), Latin, and World Language tests).

For tests with constructed-response items (Professional Readiness Examination [formerly Basic Skills]: Writing subtest, Latin, and some World Language tests), field testing also serves the purpose of obtaining examinee responses to the constructed-response items for use in training scorers.

Field test participants. Eligible participants for field testing included students enrolled in or planning to enroll in a Michigan teacher preparation program as well as candidates taking the operational test. Specific eligibility requirements for field test participants were determined depending on the field test being administered. For field tests of Professional Readiness Examination (formerly Basic Skills test) constructed-response items, all undergraduate students (freshmen, sophomores, juniors, and seniors) planning to enroll or currently enrolled in a Michigan teacher preparation program were eligible to participate. For other content-area tests, juniors or seniors enrolled in Michigan teacher preparation programs and candidates taking the operational test were eligible to participate. For the field tests for the World Language tests, juniors and seniors seeking a major or minor in the target language were also eligible to participate.

All institutions of higher education with approved teacher preparation programs were invited to have their students participate in field testing that occurred in conjunction with operational test administrations (see **Types of field testing** on the following page). In some cases, past and current MTTC examinees were invited to participate. Field test participation at stand-alone field test sessions was voluntary. Evaluation Systems developed strategies to encourage field test participation. In the case of voluntary participation, field test examinees were provided compensation in the form of gift cards or electronic gift codes, or provided with vouchers to be used for future MTTC test administration registrations.

Types of field testing. Psychometric information about test items to be used on MTTC tests was gathered through the use of field tests. Three types of field tests were conducted: operational, stand-alone at Michigan teacher preparation programs, and stand-alone at operational test administrations. These three types of field testing are described below.

*Operational field testing***.** Operational field testing is typically used to gather psychometric information on test items that have been newly developed for existing test fields. Since completion of the initial development of the MTTC in 1992, the majority of field testing has been completed on operational test forms. In this field test model, draft test items that have been reviewed and approved for use in the program are introduced onto test forms for MTTC test administrations. Items that are field tested on operational test forms are designated as non-scorable, that is, they do not contribute to the test score obtained by the examinee.

The benefits of operational field testing include the following:

- the examinee sample at an operational administration is, by definition, representative of the examinee population for that test;
- the items are administered under actual test administration conditions;
- examinees are generally motivated to perform as well as they can (since they do not know which items are field test items);
- logistical arrangements for operational field testing do not require additional activities on the part of the MDE or Michigan institutions of higher education; and
- test security provisions are stringent.

Multiple-choice items for all test fields routinely undergo operational field testing.

Stand-Alone Field Testing at Michigan teacher preparation programs. Stand-alone field testing at Michigan teacher preparation programs may be used to gather psychometric information on constructed-response items, test items for newly created test fields, and test items that would require altering standard operational administration procedures. In this field test model, draft items typically have been administered to individuals enrolled in or planning to enroll in teacher preparation programs throughout Michigan. This field test model has been used during the initial development of test fields. It has also been used to gather sample responses to the listening multiple-choice portion of the Spanish test and the constructed-response items included on the Professional Readiness Examination (formerly Basic Skills), Latin, and World Language tests.

Field test site selection and arrangements. Evaluation Systems worked with the MDE to identify potential campuses to serve as field test sites. Several factors were used in determining field test sites, including the size of the teacher preparation program, diversity in the student population at a given school, geographic location, approved endorsement area programs, and willingness of the campus to serve as a field test site. The MDE approved the list of potential field test sites that Evaluation Systems prepared. The education dean, department chair, or MTTC contact person at each field test site was contacted for approval of the use of the institution's facilities for the field test.

To make arrangements for sites that would host field test sessions, Evaluation Systems worked with a campus contact person. Evaluation Systems telephoned the contact person to describe the nature of the field test sessions, the importance of the field test to the MTTC program, and the schedule and space requirements for field testing. The role of the campus contact person was also explained, including the need to post flyers, answer questions, and receive students' registration forms, as appropriate.

After a campus was selected to be a field test site, the contact person was sent a communication acknowledging his or her agreement to participate, confirming the field test arrangements, and describing the field test administration process in greater detail.

Before the field test, Evaluation Systems worked with the campus contact to select appropriate testing rooms. Consideration was given to ensuring adequate lighting, desk space, restroom access, accessibility for all participants, controlled entrances and exits for security purposes, and the availability of parking.

Models for stand-alone field test administrations at Michigan teacher preparation programs. Two administration models were used to administer the stand-alone field tests held at teacher preparation programs: (1) intact classroom testing and (2) walk-in testing. In the intact classroom administration model, field tests were conducted during regular class periods. In the walk-in administration model, field tests were conducted in large testing rooms over the course of a day and/or evening. Field test examinees were allowed to begin the field test any time from the start of the field test session until approximately one hour prior to the close of the field test session.

Stand-alone field-testing at operational test administrations. Stand-alone field testing at operational test administrations may be used to gather psychometric information on constructed-response items, test items for newly created test fields, and test items that would require altering standard operational administration procedures. In this field test model, draft items typically have been administered to individuals currently registered for, or who have recently taken, an MTTC test in the appropriate field.

Examinee recruitment and notification. For each field test occasion, a field test flyer was developed to promote the field test at the test site and sent via e-mail to examinees and to contacts at teacher preparation programs throughout Michigan. Examinees registered for upcoming MTTC operational administrations were sent an email notifying them of the field test and inviting them to participate at the administration on the same day as their scheduled test(s). In addition, examinees who had tested during recent MTTC operational administrations or were registered to test at the operational administration, typically within the previous six months, were invited via email to participate.

Site arrangements. Field testing was held at separate, secure locations within the operational test administration sites. Examinees were screened at the entrance to the test site before proceeding to the secure field test area. The Chief Test Administrator at each field test site identified experienced test proctors (ATAs) who were trained to administer the field tests. Field testing therefore occurred in designated field test rooms at the assigned sites, proctored by dedicated ATAs. This ensured that participants responded to test items under conditions of quiet, confidentiality, and test security. Field test materials were shipped and returned with the materials for the operational test administration.

Stand-alone field test forms. The design of the field test forms depended upon the number of items to be field tested, an estimate of the likely number of field test participants, and the administration model being used. Typically, the following guidelines were used to prepare field test forms:

- test forms were created to allow field test examinees ample time to answer all test items;
- an adequate number of test forms was created to field test all draft items included in the item bank; and
- if possible, a variety of items from the item bank were included on each field test form.

Stand-alone field test procedures. The field test was administered under testing conditions approximating, to the extent feasible, an operational administration. That is, participants responded to test items under quiet conditions of confidentiality and test security. Test materials were shipped

and handled under strict security conditions. For field tests conducted at Michigan teacher preparation programs, at least one Evaluation Systems representative was present at each field test site. In addition, field test proctors were hired as needed to maintain test security. The Evaluation Systems representatives were responsible for security of test forms. They were able to answer questions related to the administration activities and to oversee the administration of the field test. For field tests conducted at operational test administrations, at least one experienced test proctor (ATA) trained to administer the field tests was present in each room designated for field testing.

Field test participants signed in to the session, showed a valid student ID (at the field tests conducted at teacher preparation programs), and were given the appropriate field test form and the necessary answer document(s). Field test sessions were monitored to ensure that participants could take the field test without undue interruption or distraction. Participants were monitored throughout the field test session to make sure that their performance represented their unaided effort. They were encouraged to make comments regarding the test items on comment forms that were part of each field test booklet.

Test security at stand-alone field tests. The security of the field test materials was maintained throughout the field test sessions. An Evaluation Systems representative or a field test proctor remained with the materials at all times. Test materials were signed out to examinees and checked when returned. Field test materials were inventoried at the end of each session to verify that all materials had been returned. Once inventoried, the field test materials were repacked for secure shipment to Evaluation Systems.

Review of field test results for multiple-choice items. The post-administration review of field test items focuses on assessing item characteristics and re-verifying item accuracy. A preliminary item analysis was conducted for each test form containing field test items. Items which demonstrated acceptable performance characteristics were retained in the item bank. Items that did not demonstrate acceptable performance characteristics were either removed from the item bank or designated as non-scorable on any future test forms.

For stand-alone field testing, Item Statistics Reports provided data on an item-by-item basis for each multiple-choice test item. The following information was included:

- percentage of examinees who answered the item correctly on the field test; and
- the item identifier indicating the objective for which the item was written.

In addition, for stand-alone field testing, a Demographic Summary Report was also prepared for each test field. The summary report listed, for each demographic question and response alternative:

- number of examinees responding to that alternative;
- unadjusted percent responding to the alternative; and
- adjusted (for no response) percent responding to the alternative.

The purpose of the Demographic Summary Report was to provide summary information that described the field test sample with respect to potentially relevant demographic variables.

Focused holistic scoring of constructed-response items. Field test responses to the constructed-response items from the Professional Readiness Examination (formerly Basic Skills), Latin, and World Language tests were scored under secure conditions, using a focused holistic scoring process. Typically, each field test response to a constructed-response item was scored independently by two scorers, with additional scoring by others as needed.

Focused holistic scoring is typically used for items intended to elicit an extended and/or complex response that must be judged according to several performance characteristics. Focused holistic scoring entails assigning scores based on an overall impression of the response, informed by considering the performance characteristics that responses are expected to display. Scores range from a low of “1” to a high of “4.” A separate code is assigned to blank or unscorable responses,

such as responses that are illegible/unintelligible, not relevant to the assigned topic, prepared primarily in a language other than the one called for by the assignment, or of insufficient length to score.

All field test responses were reviewed to check the characteristics of the draft test items and to identify potential issues that might be resolved by revision or deletion of a test item.

In addition, a summary report was prepared for the constructed-response items that were administered during field testing. The summary report listed the following information:

- distribution of scores for each constructed-response item; and
- mean score for each constructed-response item.

Rangefinding activities for constructed-response items. Rangefinding, or marker response selection, meetings were held in Michigan to review responses to the holistically scored performance assessments for the Professional Readiness Examination (formerly Basic Skills), Latin, and World Language tests and to identify marker responses that may be used in conducting operational scoring.

Members of the Content Advisory committees for Basic Skills (and the Professional Readiness Examination: Writing subtest in 2013), Latin, and World Language tests were invited to participate in the meeting(s) for their respective fields. As part of the rangefinding activities, committee members identified responses that correspond to the scoring scale points. Marker responses help define each of the score points and are used to train scorers to evaluate each examinee's performance consistently, fairly, and in accordance with the scoring scale. Marker responses represent examples of each score point, illustrating the range of responses that are found within each score point. They serve as one vehicle for calibrating future scorers to the scoring scale and may also be used as an "anchor" to maintain the scoring standard across new constructed-response items as they are developed in the future.

Marker Response Review Process. The marker response review process included the following steps for both written and oral responses:

- group training by the Chief Trainer, who described the focused holistic scoring method;
- review of the criteria used to assign a specific score to an individual response;
- review of the previously approved "1" to "4" constructed-response item scoring scale; and
- review of a subset of examinee field test responses and selection of marker responses corresponding to the score points of the scoring scale.

Chapter 5: Standard Setting

Overview

Following the first operational administration of a new or updated test, committees of Michigan educators met again to provide judgments that would assist in setting the passing standards for each test field.

The goal of standard setting was to identify standards (passing scores) for each test field that would be a fair and reasonable definition of a level of knowledge separating those endorsement candidates who had the content knowledge necessary to begin teaching acceptably in Michigan from those who did not. The standard setting process relied on professional judgments informed by input from Michigan educators who participated as content expert raters. Their judgments were provided to the MDE, who set the passing standards for the test.

The section that follows describes the procedures used from the inception of the program through October 2003. In October 2004, an updated standard-setting procedure was implemented. A description of the updated procedure follows the description of the procedures implemented through October 2003. In 2012, the question Committee members answered in providing a standard setting ratings was updated; this change is described with the section regarding procedures implemented as of October 2004.

Standard Setting Procedures through October 2003

Committee selection and notification. Content Advisory Committee members who participated in the review of test objectives and test items were invited to participate in an Item Validation and Standard Setting Conference. Each committee included school educators and teacher educators from various geographical, ethnic, and racial backgrounds, with content-specific expertise in the particular field that they reviewed.

Orientation and training. A representative of the MDE welcomed committee members, provided background information on MTTC legislation and policies, and introduced the Evaluation Systems representatives.

An Evaluation Systems representative provided background information on program activities completed to date, an overview of the conference goals, and step-by-step training for completion of the item validation and standard setting tasks. The training included:

- instructions on completing the Personal Information Form, Agreement of Security and Confidentiality, Expense Statement, and Substitute Reimbursement form;
- a description and explanation of the types of materials that were used during the conference;
- an explanation of procedures for completing the item validation and standard setting tasks; and
- a discussion of the practice exercise.

Each panel member was provided with a training manual that contained materials describing the conference procedures and review tasks. The manual included the following items.

- Descriptions of the various conference materials, including the following documents:
  - Test Framework;
  - Item Review Booklet;
  - Item Statistics Report, which included, for field test data, the sequence number, objective number, correct response, number of examinees responding, percent of examinees responding correctly (p-value), and distribution of responses of each item for which field test data were available;
  - Item Rating Form; and
  - Item Comment Form.
- Description of the item validation rating and standard setting procedures, including:
  - description (with examples) of how to rate each item with respect to validity;
  - description of how to handle special characteristics of items (i.e., typographical errors, misspelled words, illustrations that need clarification);
  - instructions on how to indicate the reason(s) for rating an item as Not Valid (relative to four validity criteria), with an example for each criterion; and
  - description (with examples) of how to rate each item with respect to standard setting.
- Practice materials

The training session included a review of the training manual. Throughout the training process, committee members were referred to specific sections of the training manual and were given opportunities to review their instructions, ask questions, and demonstrate their understanding of the procedures.

When referring to field test data, committee members were reminded that the field test data should be interpreted with caution. For example, field test data indicating the percent of examinees responding correctly reflect all examinees, not just minimally competent examinees.

Committee members participated in a practice exercise. They were instructed to read sample items in the training manual and to complete the sample rating form. Evaluation Systems staff discussed the ratings for each sample item. The sample items demonstrated different aspects of the final item validation and standard setting process.

Item validation criteria: content-area tests. Committee members used the following criteria to make the item validation ratings for content-area tests.

Objective Match

Does the item measure an important aspect of the objective?

Accuracy

Is the content of the item accurate and is the one correct or best answer designated?

Freedom from Bias

Is the item free of language or content that would disadvantage any person because of his or her gender, race, nationality, ethnicity, age, religion, handicapping condition, or cultural, economic, or geographical background?

Job-Relatedness

Does the item measure content knowledge important for the job of a Michigan entry-level educator in this content area?

Making item validation ratings. Committee members independently reviewed test items and made item validity ratings. An item was rated valid if it met the four criteria for item validity presented above: objective match, accuracy, freedom from bias, and job-relatedness. If a committee member judged an item or assignment as not valid, he or she was asked to indicate the reason(s) for considering the item invalid. Reasons for rating an item not valid included the following.

- the item does not match the objective
- the item is not accurate
- the item is not free of bias
- the item does not measure knowledge important for the job of a Michigan educator

He or she was also asked to indicate any revision(s) that would make the item valid.

Making standard setting judgments. Committee members independently reviewed test items and made standard setting judgments. The standard setting approach used was based on the procedures suggested by Angoff (1971). Committee members answered the following question for each multiple-choice item on a content-area test that he or she had rated as "valid" in the preceding step.

> "Imagine a hypothetical group of individuals who have the minimum amount of content knowledge to perform the role of a Michigan entry-level educator in this endorsement area. What percent of this group would answer this item correctly?"

Committee members used the following ten-point scale to make each standard setting rating.

1	=	0%–10%	6	=	51%–60%
2	=	11%–20%	7	=	61%–70%
3	=	21%–30%	8	=	71%–80%
4	=	31%–40%	9	=	81%–90%
5	=	41%–50%	10	=	91%–100%

Special instructions for the Basic Skills test. Members of the Basic Skills panel were asked to review test materials that are intended for all candidates, regardless of their area(s) of content specialization. Candidates for endorsements in Michigan must pass all three sections of the Basic Skills test.

Content validation. The procedures for the review of the Basic Skills test were the same as for the content area fields; the fundamental criteria for evaluating the validity of the materials were also the same. The four aspects of content validity--Objective Match, Accuracy, Freedom from Bias, and Measurement of Basic Skills Knowledge Used to Perform the Job of an Entry-Level Educator in Michigan—were to be addressed when determining the content validity of each item and prompt. Items and writing prompts that a panelist determined to meet all four criteria were to be rated "Valid."

Panelists were asked to consider the following question as the fourth validity criterion for multiple-choice items.

> Does this item measure basic skills knowledge necessary to perform successfully in Michigan teacher preparation programs?

For the writing assignments (prompts), panelists were asked to consider the following question as the fourth validity criterion.

> Does this writing assignment allow measurement of writing skills necessary to perform successfully in Michigan teacher preparation programs?

An item or prompt failing to meet any one criterion would be rated "Not Valid."

Standard setting ratings. Panelists were asked to use the following question for multiple-choice items for the Basic Skills Test.

> Imagine a hypothetical group of individuals who have the minimum amount of basic skills necessary to perform successfully in Michigan teacher preparation programs. What percent of this group would answer this question correctly?

For writing prompts, panelists were asked to review the writing assignment and the sample papers representing each of the four points on the rating scale, and answer the following question.

> Imagine a hypothetical group of individuals who have the minimum amount of writing skill necessary to perform the job of a Michigan entry-level educator. Based on the sample papers provided, which of the four points on the rating scale represents the level of writing that would be achieved by this group?

Panelists were asked to remember that the writing portion of the Basic Skills test measures general writing ability. The content or subject matter of the sample is not important so long as it could elicit a sample of writing that allows an individual to demonstrate his or her writing ability.

Item validation and standard setting results. Results of the item validation and standard setting ratings were compiled from the committee members' individual ratings. An item was considered valid if a clear majority (61% or more) of those rating the item designated it as "valid" on each of the four item rating criteria. Only items meeting this threshold could be used as scorable on MTTC test forms. Items not meeting this rule were either deleted from the item bank or used as nonscorable on MTTC test forms until they were deleted from the item bank. The committee-based preliminary minimum passing score was calculated for each test field by summing the median standard setting ratings for the scorable set of valid items. (A description of these computational procedures appears in the discussion on the next page.)

Determining passing standards. Evaluation Systems provided the MDE with the results of the standard setting activities. The MDE and STAC members reviewed the standard setting results and the Michigan Board of Education approved a final minimum passing score for the initial test

form for each test field. Revised passing scores were implemented at the appropriate test administration, as determined by the MDE.

Standard Setting Procedures beginning October 2004

In October 2004, an updated standard setting procedure was implemented. For the new tests that were made operational in October 2004 (Political Science, Biology, Integrated Science [Elementary], Integrated Science [Secondary], and Visual Arts Education), the STAC approved an extended procedure that involved additional rounds of review and decisions by the standard-setting panel. Because item validity judgments were already incorporated into the Content Advisory Committee review of items during the Item Review Conferences, the revised procedure is designed to focus on the establishment of passing standard recommendations. The steps in the revised procedure are described below.

1. Orientation and Training. Panel members were given an orientation that explained the passing score recommendation process, the materials they would be using, the concept of the "entry-level educator," and the judgments about test items and the total test that panelists would be asked to make.
2. Simulated Test-taking Activity. In order to familiarize the panel members with the knowledge and skills associated with the test items, each member was given a copy of the appropriate field's test objectives and participated in a simulated test taking experience. Each panel member was provided with a copy of the test form used for the first operational administration and was asked to read and answer the questions on the test without a key to the correct answers. After panel members completed this activity they were provided with the answer key (i.e., the correct responses to the questions on the test) and were asked to score their answers themselves.
3. Round I: Item-based Judgments. The Evaluation Systems facilitator provided training in the next step of the process, in which panel members were to make item-by-item judgments. For rating of multiple-choice items, panel members used an approach called a modified-Angoff procedure. For each item, panel members were asked to make a judgment regarding the performance of acceptably-qualified individuals.

Panel members were asked to imagine a group of Michigan educators who are just at the level of knowledge and skills required to perform the job of an entry-level educator in Michigan. Each panel member was asked to indicate what percent of this group would provide a correct response for each item. Panel members provided an independent rating for each item by answering, using their professional judgment, the following question:

> "Imagine a hypothetical group of individuals who have the minimum amount of content knowledge to perform the role of a Michigan entry-level educator in this endorsement area. What percent of this group would answer this item correctly?"

0% - 10% = 1	51% - 60% = 6
11% - 20% = 2	61% - 70% = 7
21% - 30% = 3	71% - 80% = 8
31% - 40% = 4	81% - 90% = 9
41% - 50% = 5	91% - 100% = 10

For test fields with constructed-response items, panel members made similar judgments regarding the constructed-response items on the test form they reviewed, using a procedure

known as the "extended Angoff procedure." The scoring of constructed-response items was explained to panelists. The training included a review and discussion of the performance characteristics and four-point score scales used by scorers, as well as examples of marker responses used to train scorers. The marker responses were selected to represent performance at each of the score points. Again panel members were asked to envision a group of Michigan educators who are just at the level of knowledge and skills required to perform the job of an entry-level educator in Michigan, and to provide an independent rating, from 2 to 8, to answer the following question:

> "Imagine a hypothetical individual who has the minimum amount of content knowledge to perform the role of a Michigan entry-level educator in this endorsement area. What score represents the level of response that would be achieved by this individual?"

4. Analysis of Round I Results. After the panelists completed their ratings, their rating forms were analyzed. For each test field, Item Rating Summary Reports were provided to each panelist containing for each multiple-choice item and for each constructed-response item when applicable: a) the median rating by all panelists who rated the item, b) the individual panelist's rating of the item, and c) the distribution of ratings for all panelists for the field.

5. Round II: Additional Item-Based Judgments. Evaluation Systems staff explained how to read and interpret the Item Rating Summary Reports from Round I. Panel members were also given item-level performance reports, which provided for each multiple-choice item the percent of candidates answering each item correctly at the operational test administration. Based on the additional information provided in Round II, panelists then had the opportunity to provide a second rating to replace the first round rating for each item.

6. Analysis of Round II Results. After panelists completed their ratings, the Round II rating forms were analyzed. Individual item ratings were then combined into a score that a hypothetical individual would be expected to achieve on the entire test. This recommended passing score was calculated for each panelist individually and for the group of panelists as a whole.

   The recommended passing score for each panelist was calculated by summing the panelist's individual item ratings as follows.

   For test fields containing multiple-choice items, the following steps were taken.

   A. For each scorable item, convert each rating to a number between 0 and 100 by multiplying by 10 and subtracting 5 (e.g., if the rating is 7.0, the resulting value from this step would be 65).

   B. Sum the values for all scorable items on the form and divide by 100. Round the result to the nearest integer to calculate the recommended passing score for each individual panelist.

In addition, an analogous calculation was performed for the group using the median of all panelists' ratings on each item. This recommended passing score based on item rating medians was calculated as follows.

C. Calculate the median value of all panelists' ratings for each scorable item on the test form.

D. For each scorable item, convert each median value to a number between 0 and 100 by multiplying the median by 10, subtracting 5, and rounding to the nearest integer (e.g., if the median of panelists' ratings is 7.0, the resulting value from this step would be 65; a median rating of 7.7 would result in a value of 72).

E. Sum the values for all scorable items on the form and divide by 100. Round the result to the nearest integer to calculate the recommended passing score based on item rating medians.

For test fields also containing constructed-response items, the following steps were also taken.

F. Sum the item ratings from each panel member for all constructed-response items in the section.

G. Calculate the median value of the summed ratings across all panel members.

Item-Based Passing Score Summary Reports were then distributed to panel members for each test field. For multiple-choice items, this report contained the following information:

- the number of scorable items on the test;
- the number of panelists;
- the recommended passing score for the multiple-choice section, based on item rating medians (from Step E above); and
- the distribution of individual panelists' recommended passing scores (from Step B above), sorted in descending order.

For test fields with constructed-response items, a second report was provided. This report contained the following information:

- the recommended passing score for the constructed-response section, based on item rating medians; and
- the distribution of individual panelists' recommended passing scores for the constructed-response section, sorted in descending order.

7. Test-based Passing Score Recommendation. In addition to the Item-Based Passing Score Summaries, panel members were provided with Pass Rate Analyses describing the performance of examinees at the operational test administration. Evaluation Systems staff then provided training for the test-based passing score recommendation. Panel members were instructed to consider the data in these reports and then provide a passing score (e.g., a passing score of 55 items answered correctly out of a possible 80 items on the test) in response to the following question:

> "Imagine a hypothetical individual who has the minimum amount of content knowledge to perform the role of a Michigan entry-level educator in this endorsement area. What is the number of multiple-choice items on the test that would be answered correctly by this individual?"

For test fields containing only multiple-choice items, this passing score recommendation was considered a test-based passing score recommendation. For test fields also containing constructed-response items, this passing score recommendation was considered a section-based passing score recommendation

For test fields with constructed-response items, panel members were instructed to make an additional recommendation pertaining to the constructed-response section of the test. Panel members were asked to consider the data in the reports provided and then provide a passing score in response to the following question:

> "Imagine an individual who has the minimum amount of content knowledge to perform the role of a Michigan entry-level educator in this endorsement area. What score represents the level of response that would be achieved by this individual?"

Panel members were instructed that their response to this question should be the combined total number of points out of the possible number of points for all of the constructed-response items on the test that represents the level of responses that would be achieved by this individual on the constructed-response items.

8. Determination of Passing Standards. The Standing Technical Advisory Committee (STAC) and MDE were provided with the results of the standard setting conference for review and discussion. The information provided for review and discussion are described in more detail below. The MDE then approved a final minimum passing score for each test field. For test fields containing constructed-response items, the final minimum passing score took into account both the multiple-choice and the constructed-response sections of the test. Results for all test fields are periodically reviewed by the STAC and the MDE.

Review and Implementation of Standard Setting Conference Results

Results of the standard setting recommendations are summarized for review and discussion by the MDE according to procedures approved by the STAC. First, the median of the panel recommendations for each test section are computed. Depending on the test form, there may be a single median value for an all-multiple-choice test or an all-constructed-response test, or two median values for those tests with both multiple-choice and constructed-response sections.

Along with each median, an associated standard error is computed. For the multiple-choice section, the Standard Error of the Process (SEP) is calculated. The first component of the SEP is the Standard Error of Measurement (SEM), which can be estimated form the number of multiple-choice items on the test (Lord, 1959). The estimated SEM used for the multiple-choice section of each test is:

$$\text{SEM} = 0.432 \times \sqrt{n}\text{, where } n \text{ = the number of scorable items on the test (Lord, 1959)}$$

Second, the set of judgments made by the standard setting panel have an associated Standard Error of Judgments (SEJ). The SEM and SEJ are combined according to the formula below to yield the SEP.

$$\text{SEP} = \sqrt{(\text{SEM})^2 + (\text{SEJ})^2}$$

For the constructed-response section of the test, the standard error of the panel judgments is used as the estimated standard error for the section, so for this section SEP = SEJ.

The STAC and MDE established a rule such that as each new test becomes operational, the initial passing score for each test section is set at the standard setting panel median recommendation minus 1 SEP. This -1 SEP adjustment is made to allow time for teacher candidates and teacher preparation programs to become familiar with the revised sets of test objectives and standards. The passing scores and the results from the test administrations are reviewed by the STAC annually and recommendations to adjust passing scores may be made. Any adjustments to passing scores are typically implemented at the first administration of each operational year, which usually occurs in October.

At its May 2007 meeting, the MDE and STAC agreed to a standardized procedure for adjusting the passing scores. Unless the STAC recommends otherwise, the passing scores will remain at their initial values (the panel-based median -1 SEP) for the first two years of test administration, and then will be raised to their original panel recommended values (the unadjusted panel-based median) beginning with the first test administration of the third operational year.

All revisions to the MTTC passing scores must be confirmed in writing by the MDE before they are implemented.

Standard Setting Procedures as of October 2012

In October 2012, an updated standard setting procedure was implemented. For content tests for which standards were set in October 2012 (Elementary, Early Childhood and Earth Science) and thereafter, the MDE approved an update to the standard setting question and explanation of the hypothetical candidate.

Panel members were asked to imagine a group of Michigan educators who were at the threshold of knowledge required to be an effective educator qualified to receive a Michigan teaching certificate. Panelists were instructed that the individuals used as a hypothetical reference group for their judgments should be at the level of content knowledge required to be an effective teacher in that content area.

Panelists had the following description of the hypothetical reference group for consideration in providing their ratings.

An effective educator in Michigan is expected to:

1. know and effectively teach the content defined by the test objectives;
2. effectively teach all students at a level in keeping with the high standards set for Michigan K-12 students to graduate career and college ready;
3. effectively teach all possible courses governed by the standards for the content area(s) for this certificate or endorsement; and
4. effectively teach academically advanced students as well as those who are less academically proficient within the grade levels specified by the certificate or endorsement.

Each panel member was asked to indicate what percent of this group would provide a correct response for each item. Panel members provided an independent rating for each item by answering, using their professional judgment, the following question:

> "Imagine a hypothetical group of individuals who have the level of content knowledge required to perform effectively the job of a qualified Michigan educator in this certificate/endorsement area. What percent of this group would answer this item correctly?"

Standard Setting Procedures for the Professional Readiness Examination (October 2013)

For the Professional Readiness Examination, which underwent standard setting in October 2013, the MDE approved a similarly updated standard setting question and explanation of the hypothetical candidate.

For the Professional Readiness Examination, panel members were asked to imagine a group of Individuals who have the level of knowledge and skills required to perform effectively in their student teaching assignment. Panelists were instructed that the individuals used as a hypothetical reference group for their judgments should be at the level of knowledge and skills required to be an effective student teacher.

Each panel member was asked to indicate what percent of this group would provide a correct response for each item. Panel members provided an independent rating for each multiple-choice item by answering, using their professional judgment, the following question:

> "Imagine a hypothetical group of individuals who have the level of knowledge and skills required to perform effectively in their student teaching assignment. What percent of this group would answer this item correctly?"

Panelists were also asked to provide independent judgments regarding the performance of a hypothetical candidate on each of the constructed-response assignments by answering the following question:

> "Imagine a hypothetical individual who has the level of knowledge and skills required to perform effectively in his or her student teaching assignment. What score represents the level of response that would be achieved by this individual?"

Panel members were instructed that their response to this question should be the combined total number of points out of the possible number of points for all of the constructed-response items on the test that represents the level of responses that would be achieved by this individual on the constructed-response items.

Chapter 6: Test Administration

Overview

This chapter describes policies and procedures associated with test administration activities for the MTTC, including the development of several documents related to test administration; test registration activities; the preparation of test materials before test administration; activities associated with test administration; test scoring; and data analysis and reporting.

Test Administration Planning

Initial planning documents. During the initial planning phases of the program, Evaluation Systems worked with the MDE to prepare a Test Administration Plan that provided an overview of all test administration activities and covered the following topics:

- preparation for the test administration, including the drafting of a MTTC Registration Bulletin/website update and a plan for communicating information about the update;
- registration, including the specification of services to be provided to examinees and the MDE; the deadlines for candidate registration; a description of communications to examinees submitting incomplete registration materials; storage procedures for examinee registration information; and similar matters;
- alternative testing arrangements, including procedures for examinees to request alternative testing arrangements; a plan to review each request for alternative testing arrangements on a case-by-case basis; and a specification, after case-by-case consideration, of arrangements that would be approved without further MDE review and arrangements that would require additional MDE review;
- test administration sites, staffing, materials, and procedures, as well as general parameters for dealing in an organized manner with unforeseeable situations (e.g., storms) on testing days; and
- score reporting and analyses, including preparing test materials for analysis; the types of statistical data analyses to conduct; the procedures for scoring performance assessments; and the nature of reports to be provided to examinees, institutions of higher education, and the MDE.

Additional test administration planning documents. Additional test administration planning documents provided more specific clarification of the various components of the test administrations and test scoring. Each of these documents was reviewed by the MDE. Among these additional documents were the following:

- Process for Handling Alternative Testing Arrangements for Examinees with Special Needs;
- Alternative Testing Arrangements Pre-Reviewed by the MDE;
- Michigan Test for Teacher Certification Preliminary Scoring Plan;
- Michigan Test for Teacher Certification Test Form Development and Scoring Procedures; and
- Michigan Test for Teacher Certification Guide for Information Retention (plan for retaining a reasonable level of project materials for a determined period of time).

In addition, prior to the implementation of computer-based testing in December 2010, ES provided the MDE with a Computer-Based Testing Plan that provided an overview of computer-based testing activities and covered the following topics:

- registration, including the registration process for alternative arrangements;
- test form usage and rotation;
- computer-based test administration sites, examinee admission procedures, test session information, and testing procedures; and
- scoring and score reporting, including the sore reporting timeline.

Preparation for Test Administration

Paper-based test schedule and sites. The MTTC tests have been typically offered as paper-based tests, four times a year, on Saturdays, at eleven or twelve sites throughout Michigan. As of the 2009-2010 test administration year, a fifth paper-based testing date was added. As of the 2012-2013 testing administration year, the program returned to four paper-based testing dates per year with the increasing access of computer-based testing by candidates. In addition, at one paper-based test administration date each year the MTTC is offered at two to four out-of-state sites. The MDE approves the test schedule and sites on an annual basis.

Computer-based test schedule and sites. As of December 2010, computer-based testing became available as an alternative to paper-based testing for some fields. The following table indicates the availability of computer-based testing as of fall 2015. Candidates are able to test at any one of a number of VUE Pearson Professional Centers (PPCs) and authorized test centers in Michigan, as well as at any of hundreds of PPCs throughout the United States, and in over 165 countries. Computer-based testing supplemented the four paper-based test administrations offered during the 2014-2015 program year; candidates had access to two-week testing windows in each month during which a paper-based administration is not offered. During a testing window, testing is available every day except Sunday and select holidays. The chart on the following page indicates the tests that have been made available for computer-based testing.

	12/2010	12/2011	12/2012	12/2013	11/2014	05/2015
103 Elementary Education (formerly 083)	x	x	x	Update		
196 PRE (formerly Basic Skills) - Reading	x	x	x	x		
296 PRE (formerly Basic Skills) - Mathematics	x	x	x	Update		
396 PRE (formerly Basic Skills) - Writing	x	x	x	Update		
022 Mathematics (Secondary)		x	x	x		
089 Mathematics (Elementary)		x	x	x		
002 English			x	x		
090 Language Arts (Elementary)			x	x		
106 Early Childhood Education			x	x		
009 History				x		
084 Social Studies (Secondary)				x		
093 Integrated Science (Elementary)				x		
105 Social Studies (Elementary)				x		
044 Physical Education					x	
051 School Counselor					x	
056 Cognitive Impairment					x	
059 Emotional Impairment					x	
063 Learning Disabilities					x	
017 Biology						x
018 Chemistry						x
028 Spanish						x
075 Bilingual Education						x
094 Integrated Science (Secondary)						x

Registration information and Website. Registration information for the Michigan Test for Teacher Certification is prepared for each program year, and published online on the MTTC website, with a paper version available upon request. The registration information on the website provides examinees, teacher preparation programs, and others with the specific requirements and procedures for registering to take one or more of the MTTC tests. The registration information on

the website is updated far enough in advance of the upcoming program year to provide useful and timely information about the testing program.

The Registration Information on the website typically contains the following:

- a table of contents (in the form of links on the website);
- general information about the testing program and the requirements for certification testing;
- general information about the tests, including the way they were developed and a description of the tests;
- deadlines for test registration, including regular registration, late registration, and emergency registration;
- testing fees and payment policies for both paper-based and computer-based testing;
- detailed step-by-step directions for registering to take one or more tests, for both paper-based testing and for computer-based testing;
- procedures for registering for paper- and computer-based testing;
- procedures for registering for alternative testing arrangements, for both paper-based testing and for computer-based testing;
- information about the admission tickets that properly registered examinees will receive and the notification that will be sent if registration cannot be completed;
- information on changing the test date, test site, or tests for which the examinee has registered;
- a description of what to expect at the test site, including general test site rules and the nature of the test sessions, for both paper-based testing and for computer-based testing;
- instructions to follow in the event of a potential cancellation of a paper-based test administration by the MDE;
- instructions for making comments or raising issues about the tests or the test administration or test site conditions;
- information about examinee score reports, including the process for requesting additional score reports or the cancellation of scores by the examinee;
- information about score voiding by the MTTC program;
- information about retaking a test;
- the official Rules of Test Participation governing many aspects of the test relating especially to test security, confidentiality, standardization, and fairness;
- forms needed to request test services, such as alternative testing arrangements and requesting additional score reports;
- instruction for requesting changes in registration, adding a test, or requesting a refund;
- a summary table of paper-based test administration dates and computer-based test administration testing windows, including regular, late, and emergency registration periods and deadlines for paper-based testing; and
- an estimated score report delivery date for each paper-based test administration date and each computer-based administration testing window during the period covered by the registration information.

Study guides. A study guide with preparation suggestions, test objectives, and sample test questions is available for each MTTC test. The most current study guides can be viewed or downloaded on the MTTC website.

Practice tests. The following table shows the availability of fee-based online practice tests since their inception in October 2010.

Availability of Online Practice Tests

As of Fall 2010	083 Elementary Education 096 Basic Skills (Reading, Mathematics, Writing)
Added Fall 2011	022 Mathematics (Secondary) 089 Mathematics (Elementary)
Added Fall 2012	106 Early Childhood Education (General and Special Education) 002 English 090 Language Arts (Elementary) 093 Integrated Science (Elementary)
Added Fall 2013	103 Elementary Education (replaces 083 Elementary Education) 196/296/396 Professional Readiness Examination (replaces Basic Skills) 009 History 084 Social Studies (Secondary) 105 Social Studies (Elementary)
Added Fall 2014	044 Physical Education 051 School Counselor 056 Cognitive Impairment 059 Emotional Impairment 063 Learning Disabilities

Annual revisions. Evaluation Systems updates the registration information on an annual basis, incorporating new policies and procedures as appropriate. The target date for annual updates to the website is sufficiently in advance of the first test administration in the annual cycle to give examinees adequate time to learn of the test dates and complete their registration during the regular registration period.

Registration for Paper-Based Testing

Internet registration. Internet registration is available 24 hours a day, 7 days a week, during all registration periods (i.e., regular, late, and emergency). To register for a test via the Internet, candidates go to www.mttc.nesinc.com, select "Register Now," and follow the instructions provided. Internet registration requires that candidates provide an e-mail address, social security number or Canadian Social Insurance Number, and credit card information (number, expiration date, and name of cardholder). Candidates who do not have a U.S. social security number or Canadian Social Insurance Number are instructed to request an identification number from the MDE Office of Professional Preparation Services.

In February 2010, Evaluation Systems introduced MyAccount, an interface for candidates registering on the MTTC website. MyAccount is designed to enhance candidate access to online information and registration and enables candidates to create their own password-protected

personal accounts, accessible through a unique username and password. This personal account system provides fully integrated testing services for candidates and is the single source of entry to register for all MTTC tests.

Emergency registration. Candidates who have missed the regular and late registration deadlines may register up to approximately eight days before the test administration through the emergency registration system. Through the 2013-2014 administration year, emergency registration was handled by telephone during regular business hours, Monday to Friday, excluding holidays, and by Internet 24 hours a day, 7 days a week. Beginning in October 2014, emergency registration is available via Internet only. Candidates may register, receive an examinee identification number, receive their test site and seat assignment, and make their payment by credit card through the emergency registration system.

Registration processing. Evaluation Systems processes registration applications via the Internet, or by telephone in the case of emergency registration, as they are received. Each candidate receives a unique identifying number for tracking purposes as the individual's record is changed and updated.

The information on each registration is entered into an Evaluation Systems electronic database. Until the closing date of registration, this database is updated as new registrations arrive and candidates request registration changes.

Admission ticket. Following registration, Evaluation Systems provides each candidate with a computer-generated ticket for admission to the test site for which he or she is registered. The admission ticket (one for each test) contains candidate identification information, a notice of the test(s) for which the candidate is registered, and the test site, date, and session reporting time(s) for which the candidate is scheduled.

Site rosters. Evaluation Systems prepares rosters of candidates registered at each site. The rosters, which are finalized just prior to the test date, are used to record the attendance of candidates on the test date. Information regarding the expected number of candidates is communicated to Chief Test Administrators to help them prepare for the administration.

Alternative test date for religious reasons. An alternative test date is available for candidates whose religious practices do not allow Saturday testing. Testing on the Sunday following the regular Saturday test administration date is provided.

Alternative testing arrangements because of a disability. Candidates who request alternative testing arrangements must provide adequate documentation of the disability and a recommendation from a professional regarding test accommodations. The MDE and Evaluation Systems use those recommendations to determine appropriate arrangements to be made for examinees with disabilities. The MDE and Evaluation Systems review requests for alternative testing arrangements on a case-by-case basis. The guiding principle is to provide modifications to the test and/or administration conditions that enable an equitable assessment of the required knowledge and skills that are being tested.

Registration for Computer-Based Testing

Candidates complete registration and scheduling for a computer-based test via the MTTC and Pearson VUE websites. Following a candidates' registration using MyAccount, Evaluation Systems provides candidates with access the Pearson VUE website to schedule a test appointment.

Registration is available on the Internet 24 hours a day, 7 days a week. Applicants may register up to twenty-four hours before testing. To register for a computer-based test candidates go to www.mttc.nesinc.com, select "Register Now," and follow the instructions provided. In order to register on the Internet, candidates must provide an e-mail address, social security number or Canadian Social Insurance Number, and credit card information (number, expiration date, and name of cardholder). Candidates who do not have a U.S. social security number or Canadian Social

Insurance Number are instructed to request an identification number from the MDE Office of Professional Preparation Services.

Changes and cancellations. The date, time, or location of a test appointment can be changed via the Pearson VUE website. To change a Pearson VUE confirmed appointment, the candidate must first cancel their original appointment and then schedule their new appointment. The deadline for rescheduling a test appointment is 24 hours prior to the originally scheduled testing time. There is no fee for rescheduling a test appointment by the deadline. If a test appointment has not yet been scheduled via the Pearson VUE website, candidates request withdrawal from computer-based registration by submitting a request to Evaluation Systems.

Retake policy. Candidates must wait at least 90 days before retaking a test via computer-based testing. Registering for the CBT administration does not limit registration for the paper-based test administrations of the tests available on computer.

Alternative arrangements. All Pearson Professional Centers (PPCs) are ADA compliant. Candidates who would like to request alternative arrangements should contact Evaluation Systems and complete the appropriate alternative arrangement request form and subsequent procedures before scheduling a test appointment. An Alternative Testing Arrangements Request Form for computer-based testing is provided on the MTTC website.

Once the candidate has submitted the appropriate forms and documentation, Evaluations Systems reviews the request and contacts the candidate with the determination. If the candidate's request is approved, the candidate is advised to call Pearson VUE to schedule the test appointment with alternative arrangements.

Telephone and Internet Support to Candidates

Customer Service Center. The Evaluation Systems Customer Service Center handles candidate inquiries by telephone for both paper-based and computer-based registrants through a dedicated telephone line, including a special telephone number available with a Telecommunications Device for the Deaf (TDD). Hundreds of telephone calls can be handled per hour. During peak periods (typically during the emergency registration period), additional customer service representatives are on hand to ensure efficient handling of calls.

Customer service representatives are trained in registration and administration procedures for the MTTC program and the questions that they are likely to encounter (i.e., frequently asked questions). They also receive training in telephone skills to make effective and efficient use of their encounters with candidates so that candidate concern is minimized, particularly during critical periods close to a test registration deadline. The goal is to help candidates understand the procedures of the MTTC program and to assist them in finding the information they need.

Customer service representatives have desktop access to the candidate's registration file as soon as the candidate provides identifying information. The representative displays the candidate's testing history and registration information as a way to find a helpful solution to the candidate's issues. The Customer Service Center is available by telephone during regular business hours, Monday to Friday, excluding holidays. Customer service representatives are also available during the day of the test.

Automated information system. In addition to the staff of the Customer Service Center, examinees have access 24 hours a day, 7 days a week, to the MTTC automated information system. The automated information system is a telephone-based series of recorded voice menus, activated by touch tone phone, through which the candidate may navigate a path to the information that is needed. The automated information line provides information on registration and associated deadlines, test administration, and test dates, testing fees, requests for alternative testing arrangements, emergency registration, and similar matters. In addition, candidates may request to speak with a Customer Service Center representative during regular business hours, Monday to Friday, excluding holidays.

Email communication. Candidates may also communicate with the Customer Service Center via email from the MTTC website. Most e-mail inquiries are addressed within one working day.

Preparation of Test Materials for Paper-Based Testing

Test forms. Evaluation Systems generates and sequences each test form and readies the forms for production. Identification information regarding the subarea, objective, and correct response for each item are electronically transmitted for use in data analysis. Production copies of the test forms are submitted to a final proofreading, and forms are produced in appropriate numbers to correspond to the number of candidates registered for each test at each test site, plus overage. Evaluation Systems maintains security throughout the test production process.

Test booklets. Test booklets are individually labeled with the examinee's name, identification number, seat assignment number, other identification information, and bar code labels. The booklets contain directions to the examinees and the test items.

Test administrator manuals. Evaluation Systems provides proprietary procedures manuals for Chief Test Administrators, Assistant Chief Test Administrators, and Assistant Test Administrators. The manuals are updated as needed to reflect new policies and procedures.

The manuals include information for test administrators regarding test security, receipt of test materials, test site preparation, candidate sign-in, test booklet distribution, supervision of test administrations, dealing with potential examinee collaboration, handling weather or facilities matters (e.g., power outage, flooding), returning test materials to Evaluation Systems, and similar matters.

Answer materials. Evaluation Systems provides custom-designed materials for recording responses to the test questions. Multiple-choice only tests included in the MTTC program use custom-designed answer sheets. Answer documents are used for the Professional Readiness (formerly Basic Skills test), Latin, and World Language tests. The answer documents for these tests include a section for recording responses to the multiple-choice questions and sections for recording the responses to each of the constructed-response items. Through lithocoding and bar code labels, the individual sections of the answer documents (e.g., identifying information, multiple-choice question responses, and constructed-response item responses) contain machine-readable identification information for accurate matching of different response portions for an examinee.

Approval of test materials. The test administration materials used for the MTTC, including the program website, online registration materials, test directions and instructions to examinees, and similar materials, were reviewed and approved by the MDE.

Preparation of Computer-Based Test Forms

Test forms administered on computer are established test forms that have been administered on paper and, thus, have established passing scores. As additional computer-based test forms are needed, they will be drawn from the existing set of paper-based test forms or blueprinted according to the current MTTC test form blueprinting criteria. Therefore, the same test specifications are applied to both paper- and computer-based test assembly and each computer-based test form will have undergone the equating and other quality-assurance steps associated with operational paper-based administrations. In this way, MTTC candidates receive comparable test content when testing either on paper or on computer.

Test Administration for Paper-Based Testing

Overview. Paper-based test administration for the MTTC program involves multiple steps, including identification of suitable test sites; recruitment and orientation of test administration staff; preparation and shipping of test materials under secure conditions; conducting the test administration using standardized, equitable procedures; and returning the test materials under secure conditions to Evaluation Systems.

During the 2014-2015 program year, the MTTC tests were administered four times as paper-based tests, in October, January, April, and July. The paper-based tests were offered at eleven areas throughout Michigan, and at several out-of-state locations for the April administration. Two test sessions are offered on each testing date: a four and one-half hour morning session and a four and one-half hour afternoon session. The Professional Readiness Examination (formerly Basic Skills test) is a full-session test and is administered in the morning. The World Language tests are also full-session tests; they may be administered in either the morning or the afternoon session. Examinees may register for up to two content-area tests (with the exception of World Language tests) for each test session.

Test sessions are managed by Chief Test Administrators, Assistant Chief Test Administrators, and Assistant Test Administrators, who are responsible for test security, consistency, and proper administration of the tests.

Test sites. The tests were offered in the following 11 areas during the 2014-2015 program year.

- Detroit
- Grand Rapids
- Kalamazoo
- Lansing
- Marquette
- Mt. Pleasant
- Pontiac/Auburn Hills
- Saginaw
- Sault St. Marie
- Traverse City
- Ypsilanti

Candidates who register during emergency registration may take MTTC tests in the following five areas.

- Detroit
- Grand Rapids
- Marquette
- Pontiac/Auburn Hills
- Ypsilanti

Specific test sites are selected within these areas to host tests. The areas were chosen, with MDE approval, to provide access to testing throughout Michigan.

For the April test administration only, the MTTC tests are offered at up to four out-of-state facilities, pending examinee registration.

- New York, New York
- Bloomington, Indiana
- Columbus, Ohio
- Chicago, Illinois

Site characteristics. Specifications for the MTTC test sites include these characteristics:

- A Chief Test Administrator (CTA) who meets appropriate qualifications in testing must be available. The CTA is responsible for the operation of the site including accurate and secure handling of tests, answer materials, and other related materials.
- An Assistant Chief Test Administrator (ACTA) reports to the CTA. The ACTA is responsible for assisting the CTA in the operation of the site.
- Facilities must be adequate with respect to space, furniture (desks), lighting, heat, ventilation, and accessibility for people with handicapping conditions.
- Facilities must include a limited-access secure storage area for restricted testing materials.
- The site plan must address security measures.

Testing time. The MTTC program was designed to include two four and one-half hour test sessions. The testing session is preceded by time for checking examinees in, verifying their identification, directing them to pre-assigned seats, reading directions, and permitting examinees to complete identification information in their answer documents.

In general, examinees are allowed to use the total time scheduled for the test session in the way that they prefer. Individual test sections are not subject to time limits.

Testing Personnel for Paper-based Administrations

Test administration staff for the MTTC program are responsible for ensuring consistent testing conditions and maintaining the security of the test materials. Evaluation Systems has designed procedures for the MTTC program that emphasizes control of test materials, consistency in test administration conditions, and follow-up activities in the event of suspected collaboration or other breaches in test security.

Roles and responsibilities. Testing personnel for the MTTC program include the Chief Test Administrator, the Assistant Chief Test Administrator, and a number of Assistant Test Administrators (or proctors).

The Chief Test Administrator (CTA) is ultimately responsible for pre-administration arrangements and for the implementation of the entire test administration at his or her test site. The CTA ensures that the test site is set up according to the details outlined in site specifications sent from Evaluation Systems. The CTA hires, trains, and supervises all on-site testing staff. In addition, the CTA is charged with ensuring the security and confidentiality of all testing materials and procedures. The CTA reports to the Test Administration Director at Evaluation Systems.

The Assistant Chief Test Administrator (ACTA) is responsible for assisting the CTA in implementing the test administration at his or her site. The ACTA should be able to fulfill the duties of the CTA, if necessary. The ACTA reports to and is supervised by the CTA.

The Assistant Test Administrator (ATA) is responsible for signing in up to 30 examinees and for distributing test materials to, and collecting test materials from, those examinees. The ATA also monitors the examinees and maintains test security in his or her section at all times during the test session. The ATA reports to and is supervised by the CTA.

The Evaluation Systems Representative, when present, monitors the entire testing situation and provides assistance and feedback to test site personnel as necessary.

Test administrator workshop. Chief Test Administrators (CTA) and Assistant Chief Test Administrators (ACTS) undergo comprehensive training conducted by an Evaluation Systems representative. The training is intended to provide CTAs and ACTAs with necessary information to carry out a secure and smooth test administration. CTAs are responsible for training other test administration personnel before each test administration.

Test administrator manuals. Evaluation Systems has developed and uses two proprietary test administrator manuals for the MTTC program. One manual is produced for Chief Test Administrators and Assistant Chief Test Administrators, and the other manual is produced for the Assistant Test Administrators. The manuals are designed to help ensure the uniformity of procedures across test sites. Evaluation Systems updates the manuals to accommodate changes in the program.

The manuals serve as a ready reference for test administrators. They cover the responsibilities of test administration staff and the tasks they must perform. The following information is included in the test administrator manuals:

- a description of the duties of each test administration position,
- a list of things to do in the weeks before the test administration and on the day before the administration,
- a list of the activities to be performed on the morning of the administration and during the test administration,
- a description of tasks to be performed after the test administration, and
- information about test security issues and special testing accommodations.

Throughout the testing day, administrators use the manual as a guide to accommodating examinees' needs, ensuring test site security, conducting a brief but thorough morning review of procedures with test administration staff, and supervising the entire administration.

Test administrator checklists. Evaluation Systems developed administration checklists for staff to use along with the test administrator manuals. There is one checklist for each test administration position, each with a chronological list of tasks required of that position.

Standardization. Evaluation Systems representatives periodically visit selected sites during paper-based test administrations to ensure that all personnel are following the specified procedures. Evaluation Systems representatives also evaluate the overall performance of the CTA and other test administration personnel. The CTA is responsible for training ATAs as well as for monitoring their performance. Test administration staff who do not follow proper procedures are subject to dismissal. Evaluation Systems also conducts follow-up activities with test sites and personnel after test administrations to further refine the administration process.

Packing and Shipping Test Materials for Paper-based Administrations

Before paper-based test administrations and after the registration closing date, the final number of test booklets needed is calculated, and the booklets are produced. The booklets are then gathered into sets for each Assistant Test Administrator. Each individual booklet is personalized with examinee information on the cover for each registered examinee. The booklets are packed in seat order sequence for security and ease of distribution at the test site. Support materials are gathered, including answer documents, restroom passes, and the computer-generated sign-in and inventory sheets are produced after the closing date of registration.

The materials are packed and labeled with test date, site, and ATA number. Extra booklets are packed for each session. Evaluation Systems notifies Chief Test Administrators of the number of boxes of testing materials they can expect to receive.

Candidate rosters. After the closing date of registration, Evaluation Systems prepares rosters for each site, including ATA assignments, by room, for all test sessions. The rosters, which also serve as sign-in sheets, list the candidates officially registered to take the test at a particular site. Rosters include examinees' names, specific seat assignments, identification numbers, and test code(s). Each ATA receives one candidate roster and is responsible for overseeing the administration of the test to the examinees named on the roster.

Alternative testing arrangements rosters and materials. The alternative testing arrangements approved for examinees are noted on an additional roster. The materials for examinees requiring alternative testing arrangements are shipped with the rest of the materials.

Shipping test materials. Test materials are shipped using a service that ensures acknowledgment of delivery and the traceability of materials. Back-up procedures are available for shipping, if needed.

Administration of the Paper-based Tests

On the day of the MTTC paper-based administration, test administrators make final site set-up checks and post signs for examinees. Test materials (e.g., rosters, test booklets, answer documents) are distributed to Assistant Test Administrators for the examinees who have been assigned to them.

Examinee check-in. Examinees are screened for prohibited materials (e.g., books, briefcases, calculators, and cellular phones), as stated on the MTTC website. Examinees must have their admission ticket and two pieces of identification, one with a recent photograph and signature.

Seat assignments. Examinees are directed to their assigned seats and to an Assistant Test Administrator. The ATA checks their names on the roster, makes sure they are seated properly, reviews their identification, collects their admission tickets, and has them sign the sign-in roster. Examinees whose names are not on the roster are directed to the Chief Test Administrator. Examinees are not permitted to sign-in at a test session for which they are not registered.

Overseeing testing. After examinees have been signed in and seated, the ATA distributes test materials, reads test directions, and oversees the test session, making sure that desktops remain free of inappropriate materials and monitoring examinees on an ongoing basis.

As examinees finish their tests, they signal to the ATA, who approaches them and checks and collects their materials. The ATA returns the examinee's admission ticket for use as a dismissal pass from the test site.

Michigan Department of Education participation. MDE representatives may attend paper-based test administration sessions at any and all sites during any paper-based test administration.

Return of Test Materials

At the end of each test session, ATAs log in their examinees' materials and return them to the Chief or Assistant Chief Test Administrator. The CTA, ACTA, and ATAs count the tests and answer documents to make sure that all have been returned, sign the necessary inventory sheets, and package the materials into boxes. Staff on site review and seal the boxes for return to Evaluation Systems. Both used and unused test booklets and answer documents (as well as other testing materials) are returned to Evaluation Systems using procedures similar to those used for initial distribution.

Receiving materials at Evaluation Systems. When the materials arrive at Evaluation Systems, staff recheck and count the materials in each shipment for completeness. If there are any discrepancies between the number of materials received and the number of materials indicated on the inventory sheet, Evaluation Systems follows up promptly.

Test Administration for Computer-based Testing

Overview. As of December 2010, a number of MTTC tests became available via computer-based testing, for one six-day period during each month for which paper-based testing is unavailable.

During the 2014-2015 administration year, the following ten tests were added to the list of tests available via computer-based testing:

- 044 Physical Education
- 051 School Counselor
- 056 Cognitive Impairment
- 059 Emotional Impairment
- 063 Learning Disabilities
- 017 Biology
- 018 Chemistry
- 028 Spanish
- 075 Bilingual Education
- 094 Integrated Science (Secondary)

In the winter of 2014 the number of days in a month for which each test was available by computer was expanded from six to 12 days.

Computer-based testing sites. MTTC tests administered by computer are available at a number of VUE Pearson Professional Centers (PPCs) and authorized test centers in Michigan, as well as at any of hundreds of PPCs throughout the United States, including Puerto Rico and the Virgin Islands, and in over 165 countries.

Admission to the test. To be admitted to take the test at a Pearson Professional Center, the candidate must provide one piece of current, government-issued identification printed in English, in the name in which he/she registered, and which includes a photograph and signature. All Pearson Professional Centers currently employ the following security procedures:

- Digital signature—A digital signature is required at check-in and made a part of the candidate's permanent record.
- Palm Scan or Digital photograph—A palm scan and/or a digital photograph may be taken at check-in and made a part of the candidate's permanent record.
- Access to the testing room before and after a break or on a subsequent testing session is controlled by matching the palm scan to the data recorded at initial check-in.

Test session. Candidates have four hours and 45 minutes to complete an MTTC test that would be assigned to one paper-based session. For those subject area tests for which two may be assigned in a current paper-based session, candidates are allotted two hours and 30 minutes. Included in that time is 15 minutes to read and agree to a nondisclosure agreement and to complete a brief tutorial on how to navigate the computer-based testing system. At the beginning of the test session, candidates are informed of how much time they have to complete the test. Candidates are able to monitor their time using a count-down clock that can be accessed on the screen. Candidates complete the test by selecting answers on-screen to multiple-choice items and typing responses to short-answer items and open-response items.

Once candidates complete the test session, they receive a preliminary score report for tests consisting of only multiple-choice questions or a "receipt" for tests that contain constructed-

response items that indicates they have attended the test session and instructs them to go to the MTTC website to view the computer-based testing score reporting schedule.

Chapter 7: Test Scoring and Reporting

Overview

Evaluation Systems worked with the MDE to develop scoring policies for the MTTC program. The policies include those related to scoring the Professional Readiness Examination (formerly Basic Skills) writing assignments and the constructed-response items contained in the World Language tests and the Latin test. The MTTC tests are scored under secure conditions, and test results are produced for examinees and the MDE within four weeks of the test administration date. Examinee score reports provide information to examinees and institutions of higher education that prepare teachers.

Multiple-choice Item Scoring

For paper-based testing, after examinees' answer documents are inventoried, Evaluation Systems processes and analyzes the data to produce scores and prepare score reports. The machine-scorable portion of the answer document is checked to ensure that the necessary identifying information (e.g., examinee identification information, test code, etc.) has been appropriately recorded. Each answer document is scanned and answer documents that do not complete scanning are visually checked. Examinees' responses to the multiple-choice items are electronically scored based on established answer keys.

For computer-based testing, responses to the multiple-choice questions are recorded and scored at the Pearson Professional Center and sent electronically to Evaluation Systems, where all responses are rescored for verification.

Item analyses. After each paper-based test administration, item performance is reviewed before examinee scores are reported. Items designated as scorable contribute to candidates' scores. For multiple-choice items, a preliminary item analysis is conducted for each test form. The purpose of this analysis is to verify the accuracy of the answer keys for each MTTC test. Items are flagged for review before examinees' scores are calculated. Statistics from both the current test administration and cumulative item statistics are used to identify items for review. Items are flagged if they display the following statistical characteristics:

- the percent of the examinees who answered the item correctly is less than 30 (i.e., fewer than 30 percent of examinees selected the response keyed as the correct response) ($N \geq 5$); or
- the item-to-test point-biserial correlation is less than 0.10 (if the percent of examinees who selected the correct response is fewer than 50) ($N \geq 25$);
- the modal response is not the correct response (i.e., the response chosen by the greatest number of examinees is not the response keyed as the correct response) ($N \geq 5$); or
- the percent of candidates who answered the item correctly for the most recent period decreased at least 20 points from the percent of candidates who answered the item correctly for all administrations of the item ($N \geq 25$ for the most recent period, $N \geq 50$ for all administrations).

In addition to items that are flagged due to these statistical criteria, other items may be flagged as a result of examinee comments made during or immediately after the test administration. The accuracy of each flagged item is reverified before examinees' scores are calculated.

For computer-based administrations, test forms that are were not previously administered on paper undergo the established procedures.

Constructed-response Item Scoring

The constructed-response items included on the Professional Readiness Examination (formerly Basic Skills) Writing subtest, each of the World Language tests, and the Latin test are scored under secure conditions, using a focused holistic scoring process. Typically, each response to a constructed-response item is scored by two scorers, with additional scoring by others as needed. Scorers record their scores on separate computer-scannable sheets.

Focused holistic scoring. Focused holistic scoring is typically used for items intended to elicit an extended and/or complex response that must be judged according to several performance characteristics. Focused holistic scoring entails assigning scores based on an overall, holistic impression of the response, informed by focusing on a defined set of performance characteristics associated with the scoring scale. Scores range from a low of "1" to a high of "4". A separate code is assigned to blank or unscorable responses, such as responses that are illegible/unintelligible, not relevant to the assigned topic, prepared primarily in a language other than the one called for by the assignment, or of insufficient length to score.

In the holistic scoring model, two independent scorers score examinee responses. Unless there is a discrepancy in the scores assigned by the independent scorers, the two scores are combined to determine a final score for the assignment. Discrepant scores are those that differ by more than one point. If the scores assigned are discrepant, a third scorer and/or chief scoring personnel score the response as needed to resolve the discrepancy.

Scorer selection criteria. Evaluation Systems selects scorers who are qualified according to criteria approved by the MDE. The scorer qualifications for the Professional Readiness Examination (formerly Basic Skills test) and World Language tests are presented below.

Professional Readiness Examination/Basic Skills written constructed-response item(s). To score the Professional Readiness Examination/Basic Skills written constructed-response item(s), individuals must have the following:

- English or related Language Arts degree, AND
- Teaching experience at the elementary, secondary, or college level.

Constructed-response items on the World Language tests. To score the constructed-response items on a World Language test, individuals must have the following:

- Content expertise and/or a teaching certificate in the target language, AND
- Teaching experience at the secondary or college level; OR
- College degree in the target language, AND
- Teaching experience at the secondary or college level; OR
- Fluency in the target language and a college degree.

Scoring logistics. Scoring sessions vary in duration and in the number of scorers who are asked to serve. For each scoring session, Evaluation Systems arranges for the number of scorers needed, based on the number of examinees registered for a given test administration.

Scorer orientation. Before they are permitted to score examinee responses, scorers are oriented to the background of the MTTC, the context of the scoring task, the tasks they will perform, the procedures they will follow, the scoring scale, the performance characteristics that will guide their scoring judgments, and the assignment used on test forms for the current test administration.

Typically, scorers receive practice in scoring training sets of responses to which scores have already been assigned, generally including both historic anchor responses that were used in the first test administration and current marker and training responses that apply to prompts that were used on test forms for the current test administration.

During scorer training and calibration activities, scorers must apply scores to responses in a manner that is consistent with the standards approved for the program. Performance of scorers is monitored by Evaluation Systems. Specific areas monitored during each scoring session include scorers' ability to understand and apply the established scoring scale, the consistency of the scores assigned in comparison with those assigned by the second scorer, and the scorers' ongoing consistency over time.

Only individuals who have successfully completed the orientation, including the calibration activities, are permitted to participate in the scoring session. A typical orientation session includes the following components.

The background of the program. The orientation session begins with a discussion of the testing program and its basis in Michigan laws and regulations.

The context of the scoring tasks. The scorers are given information about the context for the particular assignment that is to be scored. Scorers are informed of the part that their scoring activities play in determining the test score.

The task. Scorers receive an introduction to the fundamentals of focused holistic scoring. A basic understanding of the nature of the judgmental task is provided, in terms of its similarity with and dissimilarity from other scoring tasks that the scorers may have experienced.

Procedures. Scorers receive an orientation to the procedures that they are asked to follow, including the procedures for receiving packets of responses to score; marking scores on scoring sheets; and returning completed scoring packets to Evaluation Systems scoring staff. They are also informed of the quality control and checking procedures that are used during the scoring session.

Scoring scale. The scorers are introduced to the scoring scale(s) for the assignment(s) they will be scoring. Each constructed-response item is scored on a four-point scale, with "1" representing the low end and "4" the high end. This is followed by a discussion of each of the score points, with particular attention to the specific description of each score point.

Performance characteristics. Scorers read and discuss the performance characteristics that apply to each score point on the scoring scale. They are oriented to relating each score point on the scoring scale to the appropriate scoring criteria and performance characteristics of the responses.

The scoring trainer. The trainer, who may also be the Chief Reader, directs the orientation activities. A key aspect of training is the use of training sets of responses. Typically, training sets include responses chosen to illustrate (1) the "center" of each score point; (2) the "borderlines" between score points (e.g., high 1/low 2, high 2/low 3, high 3/low 4); and (3) "problem" responses, such as responses that are short, display poor handwriting/audio quality, or contain questionable content. Orientation also focuses on issues of equity and fair application of the scoring scale.

Scoring manual. The orientation process is facilitated by the use of several aids, including test-secure manuals to which the scorers are asked to refer during orientation and scoring activities. Scoring manuals include the following: a copy of the directions to which examinees responded; a copy of the test item(s) and prompt(s) that elicited the examinee responses; the scoring scale and performance characteristics; and a description of the procedures to be used by scorers in completing their tasks.

Calibration. Scorers must successfully complete calibration activities during the orientation process. At various points in the holistic scoring process, scorers are recalibrated to the scale, typically through discussion of specific responses. Scorers must demonstrate continued scoring accuracy on the responses; if they do not, they receive additional, individual reorientation before proceeding with scoring.

Scoring. After completing the appropriate orientation, scorers receive packets of, typically, 10 to 20 examinee responses to score. Each scorer independently reads the responses in the packet and assigns each one a score. Typically, scores are recorded on a separate, computer-scannable scoring form. As packets of responses are completed, they are turned in to scoring staff, who log them in on record-keeping sheets, remove the computer-scannable scoring forms, and recirculate them to different scorers, as appropriate.

Score reconciliation. Scoring forms for responses that have been scored twice are processed. The pair of scores assigned to each response is compared with the rules of scoring, and those responses needing further consideration are identified.

Quality control. The scoring process is designed to maintain the quality of the scores through orientation, calibration, and recalibration of scorers; monitoring of scorer performance; and Chief Reader observations.

Score Processing

After scoring is complete, results are entered into the MTTC examinee database and integrated with other examinee data, from both the other sections of the test and the registration database.

Analysis and Reporting of Results

Purposes. Evaluation Systems analyzes results of MTTC test administrations for several purposes, including:

- gathering information for equating test forms across test administrations;
- preparing reports for examinees, the MDE, and Michigan teacher preparation programs; and
- gathering statistical information on the characteristics of the tests and test items.

Test Form Equating

Overview. Equating is the process of ensuring that different forms of a test (i.e., test forms with different test items) produce similar results. Evaluation Systems uses both judgmental and statistical methods for equating test forms. Judgmental methods are used during the preparation of test forms. Statistical information is used in the production of test forms. In addition, statistical equating is completed after each test administration to ensure the comparability of different test forms.

Judgmental methods. Judgmental methods are used as test forms are created, in an effort to produce test forms that are generally similar to previous test forms in terms of certain characteristics. The characteristics that are considered as groups of potential test items are reviewed include the coverage of the test objectives in the test framework, the balance of content across subareas of the test form, the general difficulty of the items, and the general consistency from test form to test form of the test item formats that are included.

Statistical methods. Statistical methods are used to produce tests with similar statistical characteristics. While some statistical information about multiple-choice items is used as test forms are being created before test administrations, the primary application of statistical equating occurs after test forms have been composed. Statistical equating mainly focuses on the *outcomes* of the tests after they have been administered.

The central purpose of statistical equating methods is to compensate statistically for possible variability in the characteristics of test forms that may affect examinees' scores (e.g., potential chance differences in the overall difficulty level of the set of items on a new test form, compared with the overall difficulty level of the set of items on the previous test form). For example, statistical equating methods can ensure that examinees who, by chance, receive a "harder" or "easier" test form than the examinees who took the test at a previous administration must achieve a passing score that is adjusted to be equivalent to the passing score that was in effect for the previous test administration.

Equating of multiple-choice items. For the MTTC, different forms of a test are designed to contain similar numbers of test items in each subarea of the test. They also contain both scorable items (i.e., items that contribute to an examinee's score) and nonscorable items (i.e., items that are placed on a test form to gather data on their performance characteristics so that they can be used on later test forms as scorable items). Typically, test items that were formerly nonscorable are introduced onto test forms as scorable items in such a way as to produce new test forms that have performance characteristics that are similar to those of the test forms that they will replace. This process is conducted before the tests are administered to help create new test forms that will perform similarly to the test forms that they replace.

After a new test form is administered, statistical data are gathered on the multiple-choice items to verify that the new test form is similar to the previous test form. Even with the care that is taken before the test administration to achieve correspondence between the previous and new forms of the same test, the new form of a test may be slightly more or less difficult than the previous form. Test equating analyses are used after the test administration, and before examinees' scores are calculated and reported, to adjust for possible differences in test difficulty.

The equating analyses for the MTTC determine the multiple-choice item raw score (i.e., the number correct score) on the new form that is equivalent to the passing score on the previous form. A conversion equation, based on the equated passing score, is used to transform raw scores on the new test form into a set of scaled scores. The scaled scores are used for score reporting.

Equating models used for the MTTC program. Different models for selecting items for test forms (i.e., item substitution) and performing associated test score equating may be employed for the MTTC program, depending on the number of examinees who are expected to take the tests and other factors in the observed score distributions for the two test forms that are to be equated.

Z-score equating. Generally, for MTTC content area fields, each new test form is created so that all scorable items on the new form were present on the previous form as either a scorable or nonscorable item. The mean and standard deviation are computed for the set of items that were scorable on the previous test form and for the set of items that will be scorable on the new form, using the item response data of examinees who took the previous form. A score conversion method called z-score equating is applied to determine the score on the new form that corresponds to the passing score of the previous form.

The following formula is used for z-score equating:

1. Define:

   $\overline{X}$ = the mean raw score achieved on the new test form's scorable items

   s_x = the standard deviation of X

   $\overline{Y}$ = the mean raw score achieved on the previous test form's scorable items

   s_y = the standard deviation of Y

   X_i = a given raw score on the new test form

   Y_i = the raw score equivalent to X_I expressed in the raw score metric of form Y

2. Set $z_x = z_y$

$$\frac{X_i - \overline{X}}{s_x} = \frac{Y_i - \overline{Y}}{s_y}$$

3. In order to find the equivalent passing score on the new form, solve for X_i where Y_i = the passing score on the previous form.

$$X_i = \overline{X} + \left(s_x / s_y \right)\left(Y_i - \overline{Y}\right)$$

Tucker linear equating. For the Professional Readiness Examination (formerly the Basic Skills test), the set of scorable items on the new test form will be selected such that some items may be previously scorable items, some may be previously nonscorable items, and some items may be unique to the new form. For this field, the Tucker linear equating model is applied to determine the equated passing score. The set of items that is common to the previous and new forms is used to estimate total test statistics (mean and standard deviation) for scorable items on the previous form and scorable items on the new form for the total group of examinees. These statistics are then used to derive a score conversion equation to determine the score on the new form that corresponds to the passing score on the previous form.

The Tucker linear equating formula is the following:

1. Define:

   $\overline{X}$ = mean total test score on a new test form

   s_x^2 = variance of total test scores on the new form

   $\overline{U}_x$ = mean of common items on the new form

   $s_{u_x}^2$ = variance of common item scores on the new form

   s_{xu} = covariance of total test and common item set for the new form

   n_x = number of examinees taking the new form

   $\overline{Y}, s_y^2, \overline{U}_y, s_{u_y}^2$, and n_y are corresponding statistics for the previous test form

2. Compute estimated descriptive statistics for the total group (i.e., examinees from new and previous test form administrations):

$$\overline{U}_t = \left(\overline{U}_x + \overline{U}_y\right)/2$$

$$s^2_{u_t} = ((n(\overline{U}_x - \overline{U}_y)^2/2 + (n-1)(s^2_{u_x} + s^2_{u_y}))/(2n-1)$$ where n = the smaller of n_x or n_y

$$b_{xu} = s_{xu}/s^2_{u_x}$$

$$b_{yu} = s_{yu}/s^2_{u_y}$$

$$\overline{X}_t = \overline{X} + b_{xu}(\overline{U}_t - \overline{U}_x)$$

$$s^2_{x_t} = s^2_x + b^2_{xu}(s^2_{u_t} - s^2_{u_x})$$

$$\overline{Y}_t = \overline{Y} + b_{yu}(\overline{U}_t - \overline{U}_y)$$

$$s^2_{y_t} = s^2_y + b^2_{yu}(s^2_{u_t} - s^2_{u_y})$$

3. In order to find the equivalent passing score on the new form, solve for X_i, where Y_i = the passing score on the previous form.

$$X_i = \overline{X} + \left(s_x / s_y\right)\left(Y_i - \overline{Y}\right)$$

Comparability of constructed-response items. For those test with constructed-response items, assignments typically vary across test forms. Judgmental methods are used to establish the comparability of the constructed-response items that are used on successive test forms. Suitable statistical methods for equating constructed-response items are not yet widely accepted, and such methods as exist would require far greater numbers of examinees than are available in the typical administrations of the Professional Readiness Examination, Latin, or World Language tests.

The following judgmental methods of establishing the comparability of constructed-response items from test form to test form are typically used.

Scoring scales. For each constructed-response item type, an approved standardized scoring scale (with an associated set of performance characteristics) is used to assign scores to examinee responses. The scoring scale provides a written, standardized description of the "typical" response at each level of the scoring scale. The same scoring scale is used to score responses to all constructed-response items of a particular type across test administrations and across different test forms. The use of a standardized scoring scale helps to ensure the comparability of scores assigned to different individual constructed-response items within each item type.

Marker responses. Based on the score point descriptions in the scoring scale, a set of responses is selected for each constructed-response item to serve as exemplars of each point on the scoring scale. These marker responses are typically used to train and calibrate scorers to help ensure that the standardized meaning of the approved scoring scale is applied accurately and consistently to examinee responses.

Historic anchor set of marker responses. The "historic anchor set" of marker responses are usually those marker responses for the constructed-response item that was administered at the first operational administration of the test field. Future sets of responses to the constructed-response item are calibrated to this set of responses. The historic anchor set is also typically used to orient scorers for subsequent test administrations before discussing the marker responses for the current constructed-response item.

The use of the historic anchor set of marker responses, together with the standardized scoring scale and performance characteristics, helps to ensure continuity and consistency in scoring over time, across scorers, across test forms, and across test administrations. Scores retain a consistent meaning over time, and examinees' responses can be judged similarly regardless of when they take a test or which test form they take.

Scorer orientation, calibration, and monitoring. The consistency of scoring across test forms is further supported by the consistent application of standardized procedures for orienting, calibrating, and monitoring scorers during scoring sessions following each test administration. Standardized procedures for scoring are designed to help ensure that the scoring scale for each constructed-response item is applied consistently across test administrations and test forms.

Reports for Examinees, the MDE, and Institutions

Data are analyzed following paper-based test administrations and computer-based testing windows to inform examinees, the MDE, and Michigan teacher preparation programs of the passing status of individual candidates for certification. The data from each test section and test are used to prepare score reports for examinees and summary reports for the MDE that can be used in the issuance of teaching certificates. These data are also used to provide information to teacher preparation programs (and the general public) about the performance of examinees affiliated with those institutions.

Scaled Scores. The scores that are reported on the MTTC tests are "scaled" scores. They have been converted mathematically to a scale with a lower limit of 100, a passing score of 220, and an upper limit of 300.

Scaled scores and program consistency. Scaled scores provide a way to discuss consistently the different tests in the MTTC program. Because the MTTC program consists of many tests with passing scores based on varying numbers of items, it is helpful for examinees, teacher preparation programs, and the MDE to have a common metric on which to discuss test scores.

Scaled scores and equating. Scaled scores are also useful when comparing test scores from one test administration to another. The characteristics of test forms (i.e., the actual sets of test items that are administered at a given test administration) may vary slightly from one test administration to another. One of the characteristics that may vary is the combined difficulty of the test items that compose the test forms. Even though care is taken to compose test forms that are as similar as possible to test forms used previously, some variance may occur. To account for this variance and maintain an equivalent scaled passing score over time, test forms are statistically equated.

Statistical test form equating may result in changes in the number of test items within a test section that must be answered correctly to reach the level of performance represented by the state-determined passing score. For example, a total test passing score that was determined to consist, in part, of correctly answering 57 scorable multiple-choice items out of an original set of 80 scorable test items used in the passing score process may, on a set of test items that is less difficult than the original set, "equate" to correctly answering 58 scorable items. In this way, the number of test items that must be answered correctly on each test form is equated to the original number of test items that constituted the passing score.

To avoid changing the passing score that is reported to examinees, which would introduce unnecessary confusion to the score reporting process, the scaled score of 220 is uniformly used as the passing score. Thus, in the example above, correctly answering 58 scorable multiple-choice items contributes to achieving the same scaled passing score of 220 that correctly answering 57 scorable multiple-choice items contributed to previously.

Computation of scaled scores. The method used to scale scores for the multiple-choice tests and test sections in the MTTC program involves adding to 220 (the passing score) a number derived from (a) the examinee's raw score, (b) the cutscore for each test section, and (c) the maximum score that an examinee can achieve[1]. The general scaling formula is:

$$\text{scaled score} = 220 + [80 * (\text{raw} - \text{cut})/(\text{max} - \text{cut})]$$

In this formula, "raw" is the raw (observed) score on the test section. The raw score is the number of multiple-choice items answered correctly.

"Max" is the maximum possible score, the total number of points that an examinee could earn on the multiple-choice test (or test section). As with the raw score, the maximum score is expressed as a number of multiple-choice items.

"Cut" is the cutscore, the number of points determined by the state as constituting adequate performance on a test section. Again, this is expressed as a number of multiple-choice items.

Thus, "raw – cut" expresses the difference between the score the examinee actually achieved and the cutscore. If the examinee's achieved score on the test section is greater than the cutscore, "raw – cut" will be a positive number; if the examinee's achieved score on the test section is less than the cutscore, "raw – cut" will be a negative number.

"Max – cut" is the difference between the maximum possible score and the cutscore on the multiple-choice test (or test section). Since the cutscore for the tests in the program is always less than the maximum possible score, "max – cut" is always a positive number.

The fraction "(raw – cut)/(max – cut)" is calculated and is multiplied by 80. The result may be exactly 80 (if the fraction equals 1), zero or a positive number less than 80 (if the fraction equals zero or a positive number), or a negative number (if the fraction is a negative number).

If the result is 80, the examinee's scaled score for the test section will be the maximum score of 300. If the result is zero or a positive number less than 80, the examinee's scaled score for the test section will be greater than or equal to 220. If the result is a negative number, the examinee's scaled score for the test section will be less than 220. The lowest possible scaled score will vary depending on the cutscore, and can be greater than 100. Occasionally, very low performance may result in a calculated scaled score that falls below 100, in which case the lower limit is applied and the score is reported as 100.

Scaled score examples. Examples of the application of the basic scaled score formula are presented below.

The following examples assume that a test is composed of 80 scorable multiple-choice items, with a passing score of 56. Thus, max = 80 and cut = 56.

If Examinee A's raw score on the test is 80, the examinee's scaled score for the test is 300, calculated as follows:

Scaled score = 220 + [80 * (80 – 56)/(80 – 56)]

= 220 + [80 * 1/1]

= 300

If Examinee B's raw score on the test is 70, the examinee's scaled score for the test is 267, calculated as follows:

Scaled score = 220 + [80 * (70 – 56)/(80 – 56)]

= 220 + [80 * 14/24]

= 220 + 47

= 267

If Examinee C's raw score on the test is 56, the examinee's scaled score for the test is 220, calculated as follows:

Scaled score = 220 + [80 * (56 – 56)/(80 – 56)]

= 220 + [80 * 0/24]

= 220 + 0

= 220

If Examinee D's raw score on the test is 50, the examinee's scaled score for the test is 200, calculated as follows:

Scaled score = 220 + [80 * (50 – 56)/(80 – 56)]

= 220 + [80 * (-6/24)]

= 220 + (-20)

= 200

If Examinee E's raw score on the test is 20, the examinee's scaled score for the test is 100, calculated as follows:

Scaled score = 220 + [80 * (20 – 56)/(80 – 56)]

= 220 + [80 * (-36/24)]

= 220 + (-120)

= 100

Note that, for an 80-item test with a passing score of 56, an examinee with any raw score below 20 would also receive a reported scaled score of 100, because it is the lower limit of the scale. In practice, such low performance is extremely rare.

Combining the multiple-choice section and constructed-response item scores for the Writing Subtest of the Professional Readiness Examination, World Language tests (Spanish, French, and German) and Latin. The Writing Subtest of the Professional Readiness Examination, and the Spanish, French, German, and Latin tests each contain a multiple-choice section and two constructed-response items. The scaled scores of the multiple-choice section and the constructed-response item section are calculated and then combined, using weights (i.e., the percent of the total test score that is based on the test component), to produce a total test scaled score. For the Spanish, French, and German tests, the multiple-choice section accounts for 80 percent of the total test score and the constructed-response item section accounts for 20 percent of the total test score. An examinee that achieved a scaled score of 220 on the multiple-choice section and a scaled score of 260 on the constructed-response item section would receive a total test scaled score of 228, calculated as follows:

$$\begin{aligned} \text{Total test scaled score} &= (.8 * 220) + (.2 * 260) \\ &= 176 + 52 \\ &= 228 \end{aligned}$$

Similarly, for the Latin test, the multiple-choice section accounts for 90 percent of the total test score and the constructed-response item section accounts for 10 percent of the total test score. For this field, the total test scaled score would be calculated as in the example above, using weights of .9 and .1 in place of .8 and .2, respectively.

For the Writing Subtest of the Professional Readiness Examination, the multiple-choice section accounts for 50 percent of the total test score and the constructed-response item section accounts for 50 percent of the total test score. For this field, the total test scaled score would be calculated as in the example above, using weights of .5 and .5 in place of .8 and .2, respectively.

Combining the multiple-choice section and constructed-response item scores for World Language tests (Chinese [Mandarin], Arabic [Modern Standard], Russian, and Japanese). The Chinese (Mandarin), Arabic (Modern Standard), Russian, and Japanese tests each contain a multiple-choice section and eight constructed-response items comprising four sections. As with the Spanish, French, German and Latin tests, the scaled scores of the multiple-choice section and each constructed-response item section are calculated and then combined, using weights (i.e., the percent of the total test score that is based on the test component), to produce a total test scaled score. The multiple-choice section accounts for 35 percent of the total test score, the writing constructed-response item section accounts for 20 percent of the total test score, and each of the remaining three constructed-response item sections account for 15 percent of the total test score.

Writing Section of Basic Skills (through September 2013). For the writing section of the Basic Skills test, the following table is used to translate raw scores on the 2 to 8 point scale (derived from the sum of two independent scores) to scaled scores.

Raw Score	Scaled Score
2	100
3	130
4	160
5	***[2]
6	220
7	260
8	300

[2] Because all responses that receive a score of "2" from one scorer and "3" from the other scorer are considered discrepant and are resolved by additional scoring, no examinee can receive a final score of "5."

Individual score reports. The MTTC individual score reports for examinees are prepared after each paper-based test administration and each computer-based testing window. Evaluation Systems prepares individual score reports for examinees to meet a target posting date of approximately four weeks after each paper-based test administration and each computer-based testing window (except for examinees about whose tests there may be special concerns, such as a potential testing irregularity).

The MTTC score report provides the following information:

- examinee identification information;
- the name(s) of the test(s);
- the test administration date;
- notification that the examinee's scores have been sent to the Michigan Department of Education and the Michigan teacher preparation program(s), if applicable, designated by the examinee;
- the passing score for each test or subtest taken by the examinee;
- *prior to* the October 2008 test administration, the examinee's score for each test taken (for the Professional Readiness Examination/Basic Skills test, this information is provided for each subtest);
- *beginning with* the October 2008 test administration, the examinee's score for each test or subtest taken, only if the examinee did not meet the passing score;
- an indication of the status of the examinee as having met or not met the passing score for each test (or subtest for the Professional Readiness Examination/Basic Skills test);
- an indication of the examinee's performance on each subarea in each test (or Professional Readiness Examination/Basic Skills subtest) taken by the examinee; and
- for the Professional Readiness Examination/Basic Skills test, a cumulative summary of the passing status for each subtest, plus analytic scoring for the constructed-response item(s) in the writing section that provides more detailed feedback on the examinee's performance if the examinee did not pass the writing subtest.

In addition, information on how to read the score report is included with each report. This information contains information about how to read and interpret the score report, cautions relating to the use of the test score information on the report, and information about retaking a test at a subsequent test administration. Further, a "Score Report Explanation" letter may be sent to examinees who request additional information about their score reports. A sample MTTC individual score report is provided in the Appendix.

Beginning with the February 2011 computer-based testing window, Evaluation Systems introduced electronic score reporting. Examinee scores are released electronically at 10:00 p.m. eastern time on the score report dates listed on the MTTC website. Score reports are available for 45 days as PDF documents, which examinees may view, print, and download via the Score Reporting page of the MTTC website. Examinees who register online may also opt to receive their score report as an e-mail attachment. Score report e-mails are sent beginning at 5:00 p.m. on the score report date.

Post-administration summary reports. The MDE and Michigan teacher preparation programs receive reports intended to provide information regarding the performance of examinees. The institution reports summarize results for examinees who have an affiliation with the institution and, more generally, for examinees across Michigan. Following the registration period for each test administration, Michigan institutions of higher education are given the opportunity to verify that examinees who report an affiliation with their schools are eligible for inclusion on the post-administration performance summaries for their institutions. The following reports are prepared following each test administration:

- Alphabetical Score Report Roster (for each institution and statewide), and
- Performance Summary by Test Field (for eligible examinees at each institution and statewide).

All reports are made available to the Michigan Department of Education. Each institution has access through a secure website to its own reports and to files that facilitate importing examinee data for further analysis. A sample roster and summary report appear in the Appendix.

***ResultsAnalyzer*™.** In October 2011, a new reporting tool called *ResultsAnalyzer* was provided for the MDE and MTTC institutions. *ResultsAnalyzer* is a secure, password-protected, user-friendly tool that allows authorized users to view, analyze, reorganize, download, and print reports based on background information gathered during test registration and results data.

[PAGE INTENTIONALLY LEFT BLANK]

Chapter 8: Annual Program Review

Beginning with the 2014–2015 program year, an annual analysis of Differential Item Functioning (DIF) of multiple-choice items was implemented for the MTTC program as a component of bias prevention. DIF occurs when individuals of the same ability level but with different characteristics (e.g., ethnicity/gender) have different likelihoods of answering an item correctly. DIF analyses are conducted on multiple-choice items to assess whether they perform differently depending on examinees' ethnic (e.g., White/Black and White/Hispanic) or gender (male/female) group membership. Items are assessed for DIF in regard to ethnicity and gender of test takers for items with sufficient numbers of examinees in both the focal (protected) group (e.g., Hispanic) and the reference group (e.g., White). DIF analyses are generated for items on MTTC test forms that are taken by at least 100 examinees in both the focal and reference groups. Typically, and moving forward beginning with the 2015-2016 administration year, each item is assessed for DIF across all administrations of the item on each test form on which it appeared during the test administration year. For the first annual analysis, items were assessed across two program years, including data from September 1, 2013 through August 31, 2015. An item is indicated for further review if it is identified more than 50% of the time as displaying DIF, regardless of the direction of the effect (i.e., favoring the reference or the focal group).

The analysis employed the Mantel-Haenszel DIF procedure, which is designed to detect uniform DIF. Uniform DIF occurs when the probability of members from one group answering an item correctly is consistently (i.e., uniformly across all ability levels) higher than the probability of members from another group answering the same item correctly. In this procedure, examinees are sorted into focal and reference groups and matched on ability, as determined by total test score. Items were identified as differentially functioning if they met the following criterion designated by Longford, Holland and Thayer (1993): the categorization of the magnitude of DIF, represented by $|\Delta|$, is at least 1.5 and significantly greater than 1 (at the .05 significance level). Items meeting this criterion, indicating that one group performed significantly better than the group to which it was compared, were designated for further review.

Any items that flagged for DIF were designated for further review to judge if the items contained any potential bias. Items identified as differentially functioning based on gender were reviewed by the MDE, and items identified as differentially functioning based on ethnicity were reviewed by the MDE and the bias review committee (BRC). The BRC reviewed the identified items in light of the item statistics and using the standard review criteria that are applied during test development activities. Based on committee recommendations and final dispositions by the MDE, reviewed items were either retained in the banks or deleted.

[PAGE INTENTIONALLY LEFT BLANK]

Appendix: Sample Score Reports

[PAGE INTENTIONALLY LEFT BLANK]

SAMPLE MTTC INDIVIDUAL SCORE REPORT (PAGE 1 of 2)

101/001

Test Date:

See the back of this page for an explanation of how to read your score report.

SAMPLE A EXAMINEE
123 ANY STREET
SOMECITY MI 44444

Test: 083 Elementary Education **Status: Passed**

The Michigan Test for Teacher Certification (MTTC) was developed specifically for use in teacher credentialing. To preclude the use of MTTC scores for purposes other than credentialing (e.g., employment, assignment) numeric scores of passing candidates are not reported.

Number of Questions	Subarea Name	Subarea Performance Index
11 to 20	Language Arts	+++
11 to 20	Mathematics	+++
11 to 20	Social Studies	+++
11 to 20	Science	++++
1 to 10	The Arts	++
1 to 10	Health and Physical Education	+++

Test: 084 Social Studies **Total Score: 189**

Status: Did Not Pass **Minimum Passing Score: 220**

Number of Questions	Subarea Name	Subarea Performance Index
11 to 20	Historical Perspective	+
11 to 20	Geographic Perspective	++
11 to 20	Civic Perspective	+++
11 to 20	Economic Perspective	++
11 to 20	Inquiry & Public Discourse & Dec. Making	+

Examinee Name: SAMPLE A EXAMINEE Social Security Number: XXX - XX - 9999

Your results have been sent to the Michigan Department of Education and the following Michigan institutions:
University of Detroit Mercy Wayne State University Marygrove College

SAMPLE MTTC INDIVIDUAL SCORE REPORT (PAGE 2 of 2)

How to Read Your Score Report

Overview. This report provides your test results for the Michigan Test for Teacher Certification (MTTC). For each test you took, it indicates whether or not you passed, and your performance on each subarea of the test.

Test Status. Your score report indicates your status for each test you have taken. Your test status is based on all sections of the test and is reported as "Passed" or "Did Not Pass." For subject-area tests, your score for the multiple-choice test items is combined with your scores(s) on any performance assignment(s) of the test. If a test contains performance assignments (such as writing assignments), examinees must attempt all performance assignments or they cannot pass the test regardless of their scores on other subareas. The total test score is converted to a score on a scale with a range of 100 to 300. A scaled score of 220 represents the minimum passing score for each test. Your test status is listed as "Passed" if your scaled score is 220 or greater. If your scaled score is less than 220, your status is listed as "Did Not Pass" and your scaled score is provided.

For the Basic Skills, your status on each section (Reading, Mathematics, Writing) taken at this administration is provided along with your section score. In "Cumulative Status," you will find your overall status for each section: "Passed," "Not Yet Passed," or "Not Yet Taken."

Passing scores for the MTTC were established by the Michigan State Board of Education based on recommendations of committees of Michigan educators. The passing score is designed to reflect the level of knowledge and skills required for effective performance in Michigan schools. To preclude the use of MTTC scores for purposes other than credentialing (e.g., employment, assignment), numerical scores of passing MTTC candidates are not reported. The report lists your total test score only if you have not met the passing score. An indication of performance in each subarea of the test is provided for all examinees.

Subarea Performance. The report indicates your performance on each subarea of the test. This information will help you understand your areas of strength and/or weakness; you do not "pass" individual subareas.

Multiple-Choice Subarea Performance. This information should be interpreted with caution since subareas contain different numbers of questions. For each subarea containing multiple-choice questions, you will see one of the following designations:

++++ if you answered most of the questions correctly

+++ if you answered many of the questions correctly

++ if you answered some of the questions correctly

+ if you answered few or none of the questions correctly

Performance Assignments. Responses to performance assignments are reviewed and evaluated independently by at least two scorers using four-point score scales that describe varying levels of performance (the score scales are provided in the study guide, which is available on the MTTC Web site at www.mttc.nesinc.com). If you took a test with one or more performance assignments, you will see one of the following designations for each subarea containing one or more performance assignment(s):

++++ if your response(s) showed strong skills (see score point description "4")

+++ if your response(s) showed satisfactory skills (see score point description "3")

++ if your response(s) showed limited skill (see score point description "2")

+ if your response(s) showed a lack of skill (see score point description "1")

Reading a Sample Subject-Area Test Score Report. A sample score report is provided below.

Test: 09 History — **Total Score: 210** ← Ⓐ

Status: Did Not Pass — **Minimum Passing Score: 220**

Number of Questions	Subarea Name	Subarea Performance Index
31 or more	U.S. History	++
11 to 20	World History	+++ ← Ⓑ
11 to 20	Connection Among Social Science Disciplines	++ ← Ⓒ
11 to 20	Social Science Perspectives & Skills	+++

In the sample, the examinee did not pass the test (total score of 210 Ⓐ compared to the minimum passing score of 220). The examinee performed well on two subareas of the test: World History (+++) Ⓑ and Social Science Perspectives & Skills (+++). The examinee did not perform well on two subareas of the test: U.S. History (++) and Connection Among Social Science Disciplines (++) Ⓒ. When preparing to retake the test, the examinee should probably focus on the two subareas in which performance indicates weakness.

Reporting of Scores. Your scores have been forwarded to the Michigan Department of Education and to the Michigan institution(s) you indicated during registration. You should keep this score report for your own records.

Retaking the Test. If you need to retake a test, please consult the MTTC Web site, www.mttc.nesinc.com, for information on registering.

Copyright © 2008 Pearson Education, Inc. or its affiliate(s). All rights reserved.
Evaluation Systems, Pearson, P.O. Box 226, Amherst, MA 01004
Pearson and its logo are trademarks in the U.S. and/or other countries of Pearson Education, Inc. or its affiliate(s).

Rv 7/08

MTTC INSTITUTION ROSTER AND SUMMARY REPORT (PAGE 1 of 4)

Michigan Test for Teacher Certification

INSTITUTION ROSTER BY TEST

Institution: College
Test Date:

DESCRIPTION OF REPORT

This report is organized by test field. For subject area tests, score information is presented for each examinee who indicated on the registration form that s/he is currently attending your institution AND was confirmed as eligible during the institution eligibility verification process. For the Basic Skills test, score information is presented for each examinee who indicated s/he is currently attending your institution. Because the Basic Skills test is not included in the eligibility verification process, all such examinees are listed.

The results for each test, or in the case of Basic Skills each subtest, are reported on a scale ranging from 100 to 300 with the passing score set equal to a scaled score of 220. The examinee test status is liste (P=Pass; F=Not Pass; N=Not Taken) and, for each examinee who did not meet the passing score, the examinee's scaled score on the total test is provided. Following the individual examinee results is an institution sum This summary provides information for both the institution and the state as a whole regarding the number of examinees tested, the mean scaled score (scaled mean score for Basic Skills Writing), percentage of examinee passing and not passing, and the percentage of examinees with subarea performance greater than or equal to a scaled subarea score of 220. For tests for which five (5) or more examinees are reported, the mean scaled s on the test is provided for the institution. The mean scaled score computation includes the results of all reported candidates. For examinees who did not pass the Basic Skills Writing subtest, the percentage of exa in each analytic category is listed.

KEY TO SUBAREA SCORES

Each MTTC test is defined by 2 to 8 skill areas called subareas. Each Basic Skills subtest of Reading, Mathematics, and Writing is also defined by subareas. Each of these subareas has been judged important by educators in the state of Michigan and has been adopted by the Michigan State Board of Education as eligible to be tested by MTTC. For each test, or in the case of Basic Skills each subtest, the examinee's performance on each subarea is indicated. Each subarea is listed as a number, and the text for each is pres as a part of each field's report. The text may also be found in the test objectives and study guides availa on the MTTC Web site at www.mttc.nesinc.com. For the Basic Skills test, the subareas for the Reading subtes listed as 1-6, and the subareas for the Mathematics subtest are listed as 7-9. Examinee performance for each subarea is indicated as follows:

1 = Examinee answered few or none of the questions correctly (scaled subarea score of 100-179)
2 = Examinee answered some of the questions correctly (scaled subarea score of 180-219)
3 = Examinee answered many of the questions correctly (scaled subarea score of 220-259)
4 = Examinee answered most or all of the questions correctly (scaled subarea score of 260-300)

Written responses are evaluated using four-point score scales that describe varying levels of performance (these score scales are provided in the MTTC study guides). Examinee performance for each subarea is reported as follows:

1 = the response(s) showed a lack of skills (see score point description "1")
2 = the response(s) showed limited skills (see score point description "2")
3 = the response(s) showed satisfactory skills (see score point description "3")
4 = the response(s) showed strong skills (see score point description "4")

For the Basic Skills Writing subtest, a score of "0" for the written response indicates that the examinee's response was unscorable. The reason for an unscorable designation is listed as follows:

U:1 = Not on the given topic
U:2 = Illegible
U:3 = Not in the appropriate language
U:4 = Too short to score
U:5 = Blank

For examinees who did not pass the Basic Skills Writing subtest, a number corresponding to one of the seven analytic categories listed below is used to indicate each area of weakness.

1 = Appropriateness
2 = Unity and Focus
3 = Development
4 = Organization
5 = Sentence Structure
6 = Usage
7 = Mechanical Conventions

MTTC INSTITUTION ROSTER AND SUMMARY REPORT (PAGE 2 of 4)

Michigan Test for Teacher Certification

INSTITUTION ROSTER BY TEST

Institution: College
Test Date:

INTERPRETIVE CAUTIONS

1. Results reported for only a small number of examinees may not provide a valid indication of how examinees at an institution typically perform.

2. Information regarding examinee eligibility/institutional affiliation was obtained from examinees as self-reported information during the registration process. Institutions of higher education were provided with an opportunity to verify the examinees' eligibility for all subject area tests. For institutions who do not choose to participate in this verification process, and for the Basic Skills test, data included in their results are based on examinee-reported institutional affiliation.

3. Because there are differing numbers of questions for each subarea, averaging subarea information may not equal total test scores.

4. The number of questions is not necessarily divided equally among the subareas for each test.

5. Only the last five digits of the examinee's social security number are reported on this roster.

6. Individual examinee results are confidential and are to be used only for purposes of certification and entry into teacher preparation or student teaching programs. Use of the tests to compare candidates for hiring or promotion decisions is inappropriate.

*** THE INFORMATION CONTAINED HEREIN IS CONFIDENTIAL AND MUST NOT BE DISCLOSED TO UNAUTHORIZED ***
*** PERSONS. APPROPRIATE SAFEGUARDS MUST BE IMPLEMENTED BY ALL PERSONNEL TO PROTECT AGAINST IMPROPER ***
*** DISCLOSURE OF INFORMATION. ***

MTTC INSTITUTION ROSTER AND SUMMARY REPORT (PAGE 3 of 4)

Institution: College
Test 002: English

Examinee Name		SSN	Test Date	Status	Scaled Scores	Subarea Perf. 1	2	3	4	5	6	7	8	9
B	K	2222	10/17/09	P	--	2	3	4	4					
B	A	8888	10/17/09	F	212	2	2	3	3					
B	P	8888	10/17/09	P	--	3	3	3	4					
B	J	4444	10/17/09	F	202	1	2	2	3					
C	S	6666	10/17/09	P	--	4	4	4	4					
D	E	8888	10/17/09	P	--	4	4	4	4					
E	K	8888	10/17/09	F	197	1	1	3	3					
G	K	8888	10/17/09	P	--	3	4	3	4					
H	A	8888	10/17/09	P	--	3	4	3	4					
H	N	2222	10/17/09	P	--	2	3	3	4					
J	K	2222	10/17/09	F	210	2	2	2	3					
J	J	4444	10/17/09	P	--	2	2	3	4					
K	L	6666	10/17/09	P	--	3	4	4	4					
K	K	3333	10/17/09	P	--	3	2	4	4					
L	S	6666	10/17/09	F	181	1	2	2	2					
L	A	4444	10/17/09	P	--	4	3	4	4					
M	B	8888	10/17/09	P	--	2	3	2	3					
M	C	2222	10/17/09	P	--	3	4	3	3					
M	N	6666	10/17/09	P	--	3	2	3	3					
M	M	8888	10/17/09	F	212	2	2	3	3					
N	J	4444	10/17/09	P	--	3	4	3	4					
P	L	8888	10/17/09	P	--	3	4	3	3					
R	L	1111	10/17/09	P	--	3	3	3	4					
S	S	4444	10/17/09	P	--	3	3	4	4					
S	N	8888	10/17/09	P	--	2	2	3	3					
S	L	2222	10/17/09	F	215	2	2	2	3					
S	N	4444	10/17/09	P	--	3	3	3	2					
V	K	4444	10/17/09	P	--	3	3	3	3					
W	A	2222	10/17/09	P	--	3	3	3	3					

MTTC INSTITUTION ROSTER AND SUMMARY REPORT (PAGE 4 of 4)

Institution: College
Test Date:

Summary for Test 002: English

Subarea	Subarea Name	# of Items
1	Meaning and Communication	11-20
2	Literature and Understanding	21-30
3	Genre and Craft of Language	21-30
4	Skills and Processes	11-20

	Number of Examinees	Mean Scaled Score	Number (Percent) Passing	Number (Percent) Not Passing	Percentage of Examinees with Subarea Scaled Scores of 220 or Greater 1	2	3	4
Institution	29	237.2	22 (76%)	7 (24%)	59	62	83	93
Statewide	361	242.8	293 (81%)	68 (19%)	61	71	83	89

The analysis below presents summary information on performance by examinee academic major and minor, and grade level at which examinees plan to seek certification. The analysis is based upon self-reported background information gathered from examinees during the test registration process. Note that this information may be unavailable for some examinees.

	Number of Examinees	Mean Scaled Score	Number (Percent) Passing	Number (Percent) Not Passing	Percentage of Examinees with Subarea Scaled Scores of 220 or Greater 1	2	3	4
Academic Major								
Institution	13	247.2	12 (92%)	1 (8%)	77	77	92	92
Statewide	197	242.9	161 (82%)	36 (18%)	60	74	82	86
Academic Minor								
Institution	16	229.1	10 (63%)	6 (38%)	44	50	75	94
Statewide	184	242.5	132 (80%)	32 (20%)	62	68	85	93
Grade Level - Elementary								
Institution	8	226.1	5 (63%)	3 (38%)	25	50	63	88
Statewide	83	231.0	56 (67%)	27 (33%)	47	60	69	87
Grade Level - Secondary								
Institution	21	241.4	17 (81%)	4 (19%)	71	67	90	95
Statewide	273	246.5	234 (86%)	39 (14%)	66	75	88	91

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999) *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

American Psychological Association, National Council on Measurement in Education, and American Educational Research Association. (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Angoff, W. H. (1971). Norms, scales, and equivalent scores, a chapter in R. L. Thorndike (Ed.), *Educational Measurement* (2nd Ed.). Washington: American Council on Education.

Lord, F. M. (1959). Tests of the same length do have the same standard error of measurement. *Educational and Psychological Measurement*, 1959, 19, 233-239.

.

[PAGE INTENTIONALLY LEFT BLANK]

Cumulative Content Validation Survey Results

Mean Rating of Test Objectives by Test Field School Teacher and Teacher Educator Survey Results 1992-2011

Year	Field #	Field Name	# of Test Objectives	School Teacher Respondents			Teacher Educator Respondents		
				# of Respondents	Mean Rating of Test Objectives	Standard Deviation	# of Respondents	Mean Rating of Test Objectives	Standard Deviation
1992	01	Language Arts	27	109	4.12	.90	60	4.14	.93
	05	Reading	27	132	4.21	.82	54	4.24	.78
	16	Science	25	115	3.70	.97	101	3.84	.94
	22	Mathematics	31	112	3.79	1.08	84	4.04	.96
	38	Industrial Arts	26	112	3.67	.94	13	3.03	1.47
	83	Elementary Ed.	48	189	3.90	.95	130	4.10	.96
1993	09	History	29	116	3.79	.97	82	4.16	.87
	11	Psychology	28	88	3.75	.95	75	3.64	1.03
	17	Biology	31	120	3.91	.93	95	3.94	.85
	06	Social Science	27	109	3.87	1.05	89	4.07	.97
1994	82	Early Childhood Education	24	120	4.30	.83	45	4.56	.71
	02	English	19	114	4.11	.90	82	4.22	.91
	43	Health	27	109	3.75	.93	25	4.07	.85
	28	Spanish	21	118	4.03	.91	51	4.22	.87
1995	32	Business Education	26	103	3.89	1.04	23	3.93	.95
	18	Chemistry	26	124	4.02	.90	57	3.98	.93
	51	Guidance Counselor	28	123	4.18	.79	20	4.33	.86
	63	Learning Disabled	25	122	4.20	.83	14	4.48	.65
	44	Physical Education	29	117	4.03	.92	57	4.19	.86
1996	59	Emotionally Impaired	28	98	4.07	.88	17	4.58	.66
	20	Geology/Earth Science	26	111	3.98	.86	22	4.15	.85
	04	Speech	29	63	3.73	1.05	39	3.79	1.06

Year	Field #	Field Name	# of Test Objectives	School Teacher Respondents			Teacher Educator Respondents		
				# of Respondents	Mean Rating of Test Objectives	Standard Deviation	# of Respondents	Mean Rating of Test Objectives	Standard Deviation
1997	48	Library Media	23	129	4.38	.74	2	4.21	.47
	56	Mentally Impaired	26	107	4.14	.81	12	4.35	.75
	85	Middle Level	22	83	4.13	.90	15	4.42	.76
	19	Physics	24	66	3.99	.95	46	3.99	.92
1998	86	English as a Second Language	18	95	4.17	.89	7	4.33	.78
	87	Industrial Technology	26	99	3.86	.88	10	3.91	.79
	84	Social Studies	25	65	3.56	1.07	52	4.17	.86
	88	Technology and Design	18	59	3.93	.95	7	4.08	.77
1999	02	English	30	75	3.97	.91	75	4.09	.90
	40	Family and Consumer Science	24	75	4.21	.83	6	4.40	.74
	03	Journalism	18	27	3.94	1.15	3	4.67	.67
	22	Mathematics (Secondary)	18	92	4.02	.99	69	4.14	.93
	44	Physical Education	29	88	3.95	.89	49	4.19	.77
2000	91	Communication Arts (Secondary)	17	37	3.97	.85	17	4.05	1.02
	50	Computer Science	22	55	3.64	1.16	27	4.06	.99
	90	Language Arts (Elementary)	14	81	4.00	.92	39	4.23	.85
	89	Mathematics (Elementary)	18	73	3.80	1.05	58	4.09	1.01
	05	Reading	27	102	4.29	.85	22	4.36	.80
	92	Reading Specialist	30	93	4.28	.78	24	4.41	.96

Year	Field #	Field Name	# of Test Objectives	School Teacher Respondents			Teacher Educator Respondents		
				# of Respondents	Mean Rating of Test Objectives	Standard Deviation	# of Respondents	Mean Rating of Test Objectives	Standard Deviation
2001	08	Geography	22	53	4.00	.85	22	3.97	1.01
	43	Health	19	64	4.17	.81	23	4.32	.73
	09	History	27	62	3.79	1.04	59	4.19	.92
	41	Visual Arts Education	24	84	3.97	.92	36	4.16	.87
2002	10	Political Science	17	95	4.00	.87	39	4.11	.82
	17	Biology	27	100	3.94	.84	57	4.01	.85
	93	Integrated Science (Elementary)	23	92	3.77	.94	43	3.84	.91
	94	Integrated Science (Secondary)	22	93	3.96	.88	36	4.20	.76
2003	19	Physics	22	57	4.03	.97	56	4.14	.91
	18	Chemistry	22	110	3.73	1.08	60	3.92	.95
	20	Earth/Space Science	24	45	3.98	.93	37	4.03	.87
	07	Economics	16	38	3.91	.98	40	3.87	.98
2004	36	Marketing Education	18	33	4.37	.77	4	4.65	.70
	46	Dance	18	13	4.14	.88	8	4.09	.95
	97	Physical Science	17	77	4.02	.92	32	4.31	.72
	98	Business. Management, Marketing, and Technology	20	97	4.10	.83	22	4.29	.83
	99	Music	21	94	4.30	.84	40	4.47	.72
2005	901	World Languages*	14	126	4.37	.79	50	4.14	.98
	56	Cognitive Impairment	20	94	4.28	.84	10	4.48	.95
	63	Learning Disabilities	23	74	4.15	.85	29	4.37	.77

Year	Field #	Field Name	# of Test Objectives	School Teacher Respondents			Teacher Educator Respondents		
				# of Respondents	Mean Rating of Test Objectives	Standard Deviation	# of Respondents	Mean Rating of Test Objectives	Standard Deviation
2006	59	Emotional Impairment	20	68	4.20	.88	14	4.55	.78
	64	Autism Spectrum Disorder	22	97	4.51	.73	14	4.54	.86
	86	English as a Second Language	20	67	4.46	.74	13	4.60	.69
2007	26	Latin	9	15	4.31	.79	8	4.51	.69
	58	Physical or Other Health Impairment	20	112	4.35	.87	1	3.55	.76
	75	Bilingual Education	20	104	4.32	.79	9	4.82	.49
2008	04	Speech	26	120	3.93	1.01	19	4.31	.88
	296	Basic Skills (Mathematics)	13	162	3.89	.99	111	4.16	.92
2009	82	Early Childhood Education	15	151	4.61	.63	43	4.74	.55
	103	Elementary Education**	28	228	4.19	.87	112	4.40	.76
2010	007	Economics	11	39	4.08	.93	12	3.81	.89
	008	Geography	14	47	4.24	.86	11	4.30	.68
	009	History	21	121	4.19	.87	37	4.47	.72
	010	Political Science	13	114	4.18	.86	28	4.23	.96
	084	Social Studies (Secondary)	23	137	4.12	.89	28	4.33	.78
	105	Social Studies (Elementary)	22	113	4.07	.90	28	4.35	.84
2011	096	Basic Skills (Writing)	5	268	4.17	.76	107	4.42	.66
	051	School Counselor	7	161	4.32	.76	21	4.48	.71

* The World Languages framework will be used for all newly developed and redeveloped language tests.

** 103 Elementary Education was originally designed to replace the 083 Elementary Test with two separately scored tests. In 2013 the test was reconfigured as one test with 150 test questions, 120 of which are scorable. For the purposes of the Content Validation Survey, the 14 objectives contained within the framework for the two separately scored tests compiled into a single test design, and the data from the survey similarly combined.