

Standard Processes

Version 6.17.2020

New Meridian Corporation has developed the *Quality Testing Standards and Criteria for Comparability Claims* (QTS). The goal of the QTS is to provide guidance to states that are interested in including content from the New Meridian item bank and intend to make comparability claims with *other assessments* that include New Meridian content. Note that “other assessments” could mean New Meridian test forms administered in another state, or it could mean the state’s previous assessments that include New Meridian assessment content but administered or scored by a different testing vendor going forward. This document is part of a set of materials that together define a system for evaluating the types of comparability claims that can be made by participating states using New Meridian’s content.

One of the key assumptions is that states licensing New Meridian content are interested in comparing their assessment results to those of another assessment program. For example, to maintain trendlines, a state may wish to continue reporting scores on the New Meridian Scale in its newly developed assessment that include items from the New Meridian Item Bank. This is referred to as *scale score comparability*. Or, a state may be interested in comparing the percentage of student who are on track or ready for college and careers (attain Level 4 or higher) with other states, districts or schools that administer New Meridian content and use the New Meridian Level 4 cut score to determine college and career readiness (CCR)¹. This is called *readiness comparability*.

To support its intended comparability claims and provide for processes such as federal peer review, we recommend each interested state collect and submit evidence demonstrating that these types of comparisons are technically defensible. The submitted evidence is reviewed and evaluations are made about how comparable the state’s planned design, administration, and scoring processes are to the *standard processes* for the New Meridian operational test forms to determine if the desired comparisons can be supported. If the desired comparisons cannot be supported, the reviewers will provide constructive and actionable feedback on what the state needs to do to support the comparability claims.

The purpose of this document is to describe the standard processes for the New Meridian operational test forms². The document provides high-level overviews with links or references to additional documents or supporting materials published by New Meridian. Following the structure of the QTS, the information in this document is organized into three areas of a testing program that are relevant to comparability claims:

¹ “College and career readiness” (CCR) in this context refers to both being *on track* for college and careers (i.e., grade-level readiness in elementary and middle school) and *ready* for college and careers (in high school).

² New Meridian builds, as part of its annual test construction process, two types of operational test form – the *Flagship* form and the *Alternative Blueprinting Option* (or *ABO*) form. Where relevant in the document, distinctions are made between the process and/or materials for each type of form.

- Design: “What is on the test?”
- Administration: “How is the test is given?”
- Scoring: “How is test performance determined?”

The intended users of this document include states that may be interested in including New Meridian content on their assessments and would like information about the standard processes for the New Meridian forms. The document is also intended to be a companion document to the *New Meridian Comparability Review Guidelines*, which is used by independent expert reviewers to evaluate the evidence submitted by states intending to make comparability claims with their assessment that include New Meridian content.

Table of Contents

Design.....	5
Claims.....	5
Test-Taking Population	6
Assessed Content.....	6
Blueprints.....	6
Item Types.....	6
Scoring Rubrics.....	7
Testing Times and Testing Window	7
Expert Review Committees.....	8
Field Testing	8
Data Review	9
Forms Construction.....	9
Universal Design.....	9
Accommodations	9
Administration	10
Test Administration Policy and Procedures	10
Instructions for Test Administrators and Coordinators.....	10
Verbal Instructions for Test Takers	10
Printed Directions for Paper-Based Test Takers	11
Mode and Device Comparability Studies.....	11
Test Security Protocols.....	11
Data Forensics Analyses.....	11
Accessibility and Accommodation Supports.....	11
Scoring.....	13
Machine Scoring.....	13
Human Scoring.....	13
Automated Scoring	13
Measurement Model	14
Scaling and Equating	14
Disaggregated Student Group Analysis.....	15

Additional Psychometric Analysis 15
Standard Setting..... 15

Design

Claims

The claims and sub claims intended by the New Meridian forms are stated in section 2.1 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)) or available in the QTS Document Repository³:

- 1 - 2019 Technical Report – Flagship.pdf
- 1 - 2019 Technical Report – ABO.pdf

ELA/Literacy

Master Claim. The master claim is the overall performance goal for the ELA/Literacy (ELA/L) Assessment System—students must demonstrate that they are college- and career-ready or on track to readiness as demonstrated through reading and comprehending of grade-level texts of appropriate complexity and writing effectively when using and/or analyzing sources.

Major Claims.

- 1) reading and comprehending a range of sufficiently complex texts independently; and,
- 2) writing effectively when using and/or analyzing sources.

Sub Claims. The sub claims further explicate what is measured on the ELA/L assessments and include claims about student performance on the standards and evidences outlined in the evidence tables for reading and writing (see references in the [Assessed Content](#) section). The claims and evidences are grouped into the following categories.

1. Vocabulary Interpretation and Use
2. Reading Literature
3. Reading Informational Text
4. Written Expression
5. Knowledge of Language and Conventions

Mathematics

Master Claim. The degree to which a student is college- or career-ready or on track to being ready in mathematics. The student solves grade-level/course-level problems aligned to the Standards for Mathematical Content with connections to the Standards for Mathematical Practice.

Sub Claims. The sub claims further explicate what is measured on the mathematics assessments and include claims about student performance on the standards and evidences outlined in the evidence statement tables for mathematics (see references in the [Assessed Content](#) section). The claims and evidence are grouped into the following categories.

- *Subclaim A:* Major Content with Connections to Practices.
- *Subclaim B:* Additional and Supporting Content with Connections to Practices.

³ All documents in the QTS Document Repository are available upon request by emailing info@newmerdiancorp.org.

- *Subclaim C*: Highlighted Practices with Connections to Content: Expressing Mathematical Reasoning by constructing viable arguments, critiquing the reasoning of others, and/or attending to precision when making mathematical statements.
- *Subclaim D*: Highlighted Practice with Connections to Content: Modeling/Application by solving real-world problems by applying knowledge and skills articulated in the standards.

Test-Taking Population

Detailed information about the test-taking population for the New Meridian forms in 2018–19 can be found in section 11 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). The subsequent sections and Appendix 11 in the report provides more details, including demographic characteristics, of the test-taking population in 2018–19.

Assessed Content

Evidence statement are available at the following web site on the [ELA/Literacy](#) page and on the [Mathematics](#) page or available in the QTS Document Repository:

- 5a - Reading Evidence Statements.pdf
- 5b - Writing Evidence Statements.pdf
- 5c - Mathematics Evidence Statements.pdf

Description about evidence statements from the web page: “Evidence Statements describe the knowledge and skills that an assessment item or a task elicits from students. They are aligned directly to the Common Core State Standards and highlight their advances especially around the coherent nature of the standards.”

Blueprints

Blueprints for the New Meridian flagship forms are available at [ELA/L Grades 3-5](#), [ELA/L Grades 6-11](#), and [Mathematics](#). Blueprints for the New Meridian ABO forms are available at [ELA/L](#) and [Mathematics](#). All blueprint documents are also available in the QTS Document Repository:

- 4a - ELA Literacy Grade_3-5 Blueprint (Flagship).pdf
- 4b - ELA Literacy Grade_6-11 Blueprint (Flagship).pdf
- 4c - Mathematics Blueprint (Flagship).pdf
- 4d - ELA Literacy Blueprint (ABO).pdf
- 4e - Mathematics Blueprint (ABO).pdf

Item Types

Item types on the New Meridian forms include:

- Multiple choice
- Multiple select
- Drag and drop
- Inline choice

ELA/L Only

- Evidence Based Selected Response (EBSR)
- Text extraction
- Text highlighting
- Match table grid
- Constructed response including Prose Constructed Response (PCR)

Mathematics Only

- Equation editor
- Fill-in-the-blank
- Hot spot
- Drawing tool
- Coordinate plane
- Line graph
- Bar graph
- Function graphs
- Histogram
- Interactive number line
- System of equations
- Graphing polygons
- Composite Graph
- Fraction Model
- Zoom Number Line

Scoring Rubrics

The scoring rubrics for prose constructed response (PCR) writing tasks in the New Meridian item bank can be found on the [ELA/Literacy page](#), or from the QTS Document Repository:

- 3 - Grades 3-11 ELA Literacy Scoring Rubric.pdf

Testing Times and Testing Window

The unit testing times for the New Meridian forms are summarized in the tables below.

Spring 2019 Unit Testing Times for Grades 3-5

Subject(s)	Unit	Flagship Testing Time (Minutes)	ABO Testing Time (Minutes)
Mathematics Grade 3 Mathematics Grade 4 Mathematics Grade 5	Unit 1	60	60
	Unit 2	60	60
	Unit 3	60	60
	Unit 4	60	NA
ELA/Literacy Grade 3	Unit 1	75	75
	Unit 2	75	75
	Unit 3	75	NA

Subject(s)	Unit	Flagship Testing Time (Minutes)	ABO Testing Time (Minutes)
ELA/Literacy Grade 4 ELA/Literacy Grade 5	Unit 1	90	90
	Unit 2	90	90
	Unit 3	90	NA

Spring 2019 Unit Testing Times for Grades 6-8

Subject(s)	Unit	Flagship Testing Time (Minutes)	ABO Testing Time (Minutes)
Mathematics Grade 6	Unit 1	80	60
Mathematics Grade 7	Unit 2	80	60
Mathematics Grade 8	Unit 3	80	60
ELA/Literacy Grade 6	Unit 1	90	90
ELA/Literacy Grade 7	Unit 2	90	90
ELA/Literacy Grade 8	Unit 3	90	NA

Spring 2019 Unit Testing Times for High School

Subject(s)	Unit	Flagship Testing Time (Minutes)	ABO Testing Time (Minutes)
Algebra I, Geometry, Algebra II, Integrated Mathematics I, II, III	Unit 1	90	90
	Unit 2	90	90
	Unit 3	90	NA
ELA/Literacy Grade 9	Unit 1	90	90
ELA/Literacy Grade 10	Unit 2	90	90
ELA/Literacy Grade 11	Unit 3	90	NA

Expert Review Committees

Descriptions of the expert review committees involved in the New Meridian item and test development process are provided in sections 2.2.2 and 2.2.3 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). The committee includes:

- Text Review Committee
- Content Item Review Committee
- Bias and Sensitivity Item Committee
- Editorial Review Committee
- Data Review Committee
- Test Construction Committee

Field Testing

Currently, the New Meridian forms use an *embedded census field-testing model* that includes all students who take the summative test forms for mathematics, and a *sampling approach in which a pre-selected group of students take an additional field test section* as part of their operational test form for

ELA/L. More information about the field testing model are in section 2.2.5 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

Data Review

An overview of the data review process for items in the New Meridian bank is in section 2.2.2 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

Forms Construction

An overview of the operational forms construction process for the New Meridian forms is given in section 2.2.3 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

Universal Design

The universal design requirements for item development are listed in the *Accessibility Features and Accommodations* document, which includes information about tools and accessibility features available to all students taking the New Meridian forms. The document is available for download on the [Resources](#) page, or in the QTS Document Repository:

- 9 - Accessibility Features and Accommodations.pdf

Accommodations

Guidelines, principles, and general participation policies for accommodations on the New Meridian forms are described in the *Accessibility Features and Accommodation* document. Information about the accommodations available for students with disabilities are in section 3 (*Accommodations for Students with Disabilities* subsection). The document is available for download on the [Resources](#) page, or in the QTS Document Repository:

- 9 - Accessibility Features and Accommodations.pdf

Language Translation

New Meridian offers a paper-based version of the mathematics assessment in Spanish, and both large print and Text-to-Speech versions of the test in Spanish.

Information about the translation (or transadaption) process for the Spanish forms is available in section 2.2.3 (*Spanish-Language Assessments for Mathematics* subsection) of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

Administration

Test Administration Policy and Procedures

Test administration policy and procedures can be found in the following documents in the QTS Document Repository:

- Computer-Based Testing Test Administrator Manual (Section 2.0):
 - 10a - Flagship Test Administrator Manual – CBT.pdf
 - 10d - ABO Test Administrator Manual – CBT.pdf
- Paper-Based Testing Test Administrator Manual (Section 2.0):
 - 10b - Flagship Test Administrator Manual – PBT.pdf
 - 10e - ABO Test Administrator Manual – PBT.pdf
- Test Coordinator Manual (Section 2.0):
 - 10c - Flagship Test Coordinator Manual.pdf
 - 10f - ABO Test Coordinator Manual.pdf

Instructions for Test Administrators and Coordinators

Instructions for administering the New Meridian operational test forms before, during and after testing are specified in the following documents in the QTS Document Repository:

- Computer-Based Testing Test Administrator Manual (Sections 3.0 to 5.0):
 - 10a - Flagship Test Administrator Manual – CBT.pdf
 - 10d - ABO Test Administrator Manual – CBT.pdf
- Paper-Based Testing Test Administrator Manual (Sections 3.0 to 5.0):
 - 10b - Flagship Test Administrator Manual – PBT.pdf
 - 10e - ABO Test Administrator Manual – PBT.pdf
- Test Coordinator Manual (Sections 3.0 to 5.0):
 - 10c - Flagship Test Coordinator Manual.pdf
 - 10f - ABO Test Coordinator Manual.pdf

Verbal Instructions for Test Takers

Scripts with the verbal directions that test administrators should give to test takers can be found in the following documents in the QTS Document Repository:

- Computer-Based Testing Test Administrator Manual, Section 4.11 for mathematics and Section 4.12 for English language arts/literacy:
 - 10a - Flagship Test Administrator Manual – CBT.pdf
 - 10d - ABO Test Administrator Manual – CBT.pdf
- Paper-Based Testing Test Administrator Manual, Section 4.11 for mathematics and Section 4.12 for English language arts/literacy:
 - 10b - Flagship Test Administrator Manual – PBT.pdf
 - 10e - ABO Test Administrator Manual – PBT.pdf

Printed Directions for Paper-Based Test Takers

Examples of the printable paper-based tutorials are available in the following documents in the QTS Document Repository:

- 10b - Flagship Test Administrator Manual – PBT.pdf
- 10e - ABO Test Administrator Manual – PBT.pdf

Mode and Device Comparability Studies

Mode (computer-based tests vs. paper-based tests) and device (desktop/laptops vs. tablets) comparability studies have been conducted on the New Meridian forms. High-level descriptions of the design and findings of these studies are available in section 14.5.4 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

If more detailed information about the studies is desired, it can be requested by emailing New Meridian at info@newmerdiancorp.org.

Test Security Protocols

The test security policies and procedures, including protocols for maintaining the security of test materials and content and handling testing irregularities and security breaches, are described in the following documents in the QTS Document Repository:

- Computer-Based Testing Test Administrator Manual (Sections 2.1 and 2.2): 10a - Spring 2019 Computer-Based Test Administration Manual.pdf.
- Paper-Based Testing Test Administrator Manual (Sections 2.1 to 2.2): 10b - Spring 2019 Paper-Based Test Administration Manual.pdf.
- Test Coordinator Manual (Sections 2.1 to 2.2): 10c - Spring 2019 Test Coordinator Manual.pdf.
- Security Policies and Administration Procedures Q&A: 12 –Security Policies and Administration Procedures Q&A.pdf.

Data Forensics Analyses

The 2019 Technical Report ([Flagship version](#), [ABO version](#)) includes overviews of the data forensics analysis methods in section 3.4. The analysis methods include:

- Response change analysis;
- Aberrant response analysis;
- Plagiarism analysis;
- Longitudinal performance monitoring;
- Internet and social media monitoring; and,
- Off-hours testing monitoring.

Accessibility and Accommodation Supports

Guidelines and directions for administering accessibility features and accommodations on the New Meridian forms are in the following documents in the QTS Document Repository:

- Accessibility Features and Accommodation Manual (Section 3 and the Appendices): 9 - Accessibility Features and Accommodations Manual (6th Edition).pdf
- Computer-Based Testing Test Administrator Manual (Section 6.0): 10a - Spring 2019 Computer-Based Test Administration Manual.pdf.
- Paper-Based Testing Test Administrator Manual (Sections 6.0): 10b - Spring 2019 Paper-Based Test Administration Manual.pdf.
- Test Coordinator Manual (Sections 2.1 to 2.2): 10c - Spring 2019 Test Coordinator Manual.pdf.
- Accessibility Features and Accommodations FAQ: 13 - Accessibility Features and Accommodations FAQ.pdf.

Scoring

Machine Scoring

Descriptions of the review process for key-based and rule-based machine-scored items are in section 4.1 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

A *test suite* is a set of test cases used to validate that the scoring rules and scanning rules and procedures for machine-scored items are being applied correctly. A test suite for New Meridian bank assets should include simulated responses to items in each content area (ELA/L and mathematics) and language version (English or Spanish) of the mathematics test. The test cases can contain examples of all relevant responses to items with complex scoring rules, unusual response patterns (e.g., multiple responses to a single item, blanks), variations on responses to gridded response mathematics items, as well as partially completed and partially erased responses.

Human Scoring

Section 4.2 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)) provides details of the process for human scoring (also referred to as handscoring). It includes information about:

- Scorer training and qualification (sections 4.2.1 and 4.2.2);
- Double scoring (section 4.2.4) for PCR items; and,
- Scorer monitoring, including backreading, validity sets, and the use of calibration responses and inter-rater agreement statistics for scorer intervention (section 4.2.4).

Automated Scoring

Intelligent Essay Assessor (IEA)

Currently, the automated scoring engine, Intelligent Essay Assessor (IEA), is used to score responses for PCR tasks on the New Meridian CBT (online) forms. Information and research-based evidence for the validity and reliability of scores from the IEA are available in the following documents:

- IEA fact sheet – available in the QTS Document Repository (14a - Intelligent Essay Assessor (IEA) Fact Sheet.pdf.)
- IEA frequently asked questions (FAQ) – available in the QTS Document Repository (14b - Intelligent Essay Assessor (IEA) FAQ.pdf.)
- IEA research paper – available in the QTS Document Repository (14c - Intelligent Essay Assessor (IEA) Research Paper.pdf.)

Automated Scoring Process

Section 4.3 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)) provides details of the automated scoring process for PCR tasks. It includes information about:

- Training IEA's automated scoring engine
- Criteria for evaluating the quality of scores from IEA
- Rules for assigning final scores to PCR responses that include scores from IEA
- Analysis conducted to compare scoring quality for various demographic student groups

Measurement Model

The two-parameter logistic/generalized partial credit model (2PL/GPC: Muraki, 1992) is the measurement model that underlies the New Meridian item bank. 2019 Technical Report ([Flagship version](#), [ABO version](#)) provides the following information about the measurement model:

- A description of 2PL/GPC model is given in section 7.3.1
- The calibration software, IRTPRO for Windows (version 4.2), in section 7.3.4. Additional information about IRTPRO is available at the [Vector Psychometric Group](#) web page.
- A description of the calibration procedures, including item exclusion rules, and convergence criteria is provided in sections 7.3.3 and 7.3.4.
- A description of the approach for calibrating PCR items to account for the local dependency of traits is given in section 7.3.2.

Scaling and Equating

Reporting Scales

Information about the New Meridian reporting scales are provided in section 12.2 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). This includes:

- The summative score scale: 650-850 (with 750 representing Level 4, college- and career-readiness),
- The ELA/L reading and writing claim scale: 10 to 90 for reading and 10 to 60 for writing, and,
- The sub claim scale: reported as one of three performance levels (below expectations, nearly meets expectations, meets or exceeds expectations).

Score Conversion Tables

A description of the process for generating score conversion tables for the New Meridian forms is given in section 12.3 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). As a high-level summary:

1. An IRT inverse test characteristic curve (TCC) approach is used to develop the relationship between raw scores and IRT ability estimates (θ).
2. Scale transformation constants (see Appendix 12.2) are used to transform θ s to the reporting scale.

Linking Design

The linking design for the New Meridian operational test forms is described in section 2.2.4 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). Below is an excerpt that provides an overview of the linking design:

“To support the goal of score comparability within and across administrations and years, a hybrid approach was implemented that incorporated the strengths of common item linking and randomly equivalent groups. The use of repeated operational core items was leveraged for common item linking. In addition, all forms were available throughout the operational administration, with spiraling at the student level, leveraged to support linking through randomly equivalent groups.

The operational test forms involved various types of linking; horizontal linking and across-administration linking. Horizontal linking consisted of linking items, or common items, included in both forms in a single administration. Across-administration linking, or year-to-year linking, consisted of common items included in two different administrations. The placement of linking items across forms or administrations supports the development of comparable scores.”

Equating Procedures

Section 7.6 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)) describes the year-to-year linking method based on the Stocking and Lord (1983) test characteristic curve scale transformation method. Sections 7.7 describe rules for excluding items from the linking set.

Quality Control

The quality control procedures for the item calibration, and scaling and equating of the New Meridian forms are outlined in sections 7.3.5 and 7.6 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). More details about the quality control processes can be found in section 14.7.

Disaggregated Student Group Analysis

A description of the differential item functioning (DIF) analysis based on disaggregated student group performance, conducted annually on items in the New Meridian bank, is given in sections 6.2 and 6.3 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)). These sections include information about:

- DIF procedures
- DIF classification criteria
- DIF comparison groups
- Sample size requirements for DIF

The process for evaluating items for potential exclusion from score reporting is summarized in section 7.5.

Additional Psychometric Analysis

Reliability analysis and validity studies conducted on the New Meridian forms are summarized in sections 13 and 14 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

Standard Setting

College and Career Readiness

Information about New Meridian’s definition of “college and career ready” can be found in [this document](#) or in the QTS Document Repository:

- 15a - CCR Determination Policy and PLDs_FINAL.pdf

Performance Level Descriptors

The policy claims (or policy definitions) for the New Meridian forms are articulated in section 8.2 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

The grade- and subject-specific PLDs are available for download at the [ELA/Literacy](#) and [Mathematics](#) web pages or available in the QTS Document Repository:

- 15b - ELA PLDs_All Grades.pdf
- 15c - Mathematics PLDs_All Grades.pdf

Standard Setting Process

Details of the standard setting process for the New Meridian performance levels are given in section 8.3 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).

Standard Setting Validity Studies

Examples of research studies conducted to validate its cut scores include the *Benchmarking Study*, the *Postsecondary Educators' Judgment Study*, and the *College and Career Readiness Benchmark Validation Study*.

These studies are available at the [Education Resources Information Center \(ERIC\)](#) web page or can be requested by emailing New Meridian at info@newmerdiancorp.org.

Evidence of valid interpretation and use of the New Meridian performance standards is provided in sections 14.7 and 14.8 of the 2019 Technical Report ([Flagship version](#), [ABO version](#)).