



## Virginia Department of Education

RFP for Student Growth Assessments (# DOE-SGA-2012-15)

October 29, 2012

Technical Proposal - Redacted Copy



NWEA™

Northwest Evaluation Association

*Partnering to help all kids learn*

October 23, 2012

Ms. Monique Robinson  
Virginia Department of Education  
James Monroe Building  
Procurement Office  
101 N. 14<sup>th</sup> Street, 21<sup>st</sup> Floor  
Richmond, VA 23219

Dear Ms. Robinson:

Northwest Evaluation Association™ (NWEA™) is pleased to respond to the Virginia Department of Education RFP for Student Growth Assessments (RFP# DOE-SGA-2012-15). With more than thirty-five years of experience in measuring individual student growth, and with the adaptive nature of our Measures of Academic Progress® (MAP®) assessments, we believe NWEA can offer the divisions of Virginia assessment systems that are uniquely capable of helping achieve the Commonwealth's goals.

NWEA's assessment system provides:

- Highly accurate measurement of individual student performance and growth, using a criterion-referenced assessment that is aligned with the content in the Virginia Standards of Learning.
- A very stable, vertically aligned, linear scale for scoring. Our assessments provide an accurate determination of instructional level and academic growth history for each student.
- An adaptive assessment experience, with a range of items available for each student. Division staff will receive more accurate and more complete data about each student, and particularly about those students who are performing significantly above or below their grade levels.
- Individual student results displayed immediately after completion of the assessments.
- Reporting tools that are linked to a number of resources, including practical classroom resources and ongoing professional development opportunities.

NWEA is excited about the prospect of a partnership with the divisions of the Commonwealth of Virginia to better serve children. We hope this proposal will lead to an invitation to meet, share ideas, and create a solution that will address the needs of the children in your school system.

Best regards,



Jeffrey P. Strickler  
Executive Vice President and COO

## Table of Contents

RFP Form .....	4
Addenda 1 .....	6
Pricing Schedule .....	8
Proposal Narrative.....	10
Overview .....	10
a. Organizational Structure.....	13
b. Maintaining Confidentiality of Student Data .....	14
c. NWEA’s Training Plan .....	15
d. Using MAP Assessments as Measures of Growth.....	19
e. Demonstrated Experience .....	20
f. NWEA’s References.....	22
Attachment 1 – Small Business Subcontracting Plan .....	24
Attachment 2 – State Corporation Commission Form.....	28
Attachment 3 – Assessment Description Template .....	30
Attachment 4 – Assessment Description Template .....	31
Attachment 5 – Assessment Description Template .....	32
Copies of Proposed Assessment .....	110
Instructions for Accessing Sample Administrative and Student Site .....	110
Continuity of Services.....	116



NWEA™

Northwest Evaluation Association

*Partnering to help all kids learn®*

(This page left intentionally blank)

**COMMONWEALTH OF VIRGINIA  
DEPARTMENT OF EDUCATION  
REQUEST FOR PROPOSAL (RFP)**

Issue Date: September 25, 2012 RFP# DOE-SGA-2012-15  
Title: Student Growth Assessments  
Commodity Code: 92420  
Issuing Agency: Commonwealth of Virginia  
Virginia Department of Education  
101 North 14<sup>th</sup> Street, 21<sup>st</sup> Floor  
Richmond, Virginia 23219  
Using Agencies: Local Education Agencies including  
Virginia Public School Divisions and  
Virginia Public Schools  
Work to be Performed: Offsite

Initial Period Of Contract: From Date of Award Through November 15, 2014; (Renewable).

Sealed Proposals Will Be Received Until 2:00 PM October 29, 2012, For Furnishing The Goods/Services Described Herein.

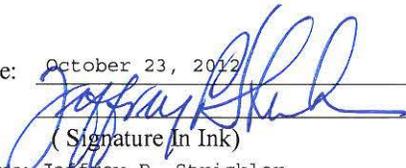
All Inquiries For Information Should Be Directed To: Marie Williams, Contract Officer, at (804) 225-2040.

PROPOSALS MUST BE DELIVERED TO THE JAMES MONROE BUILDING, 101 N. 14<sup>TH</sup> STREET, RICHMOND, VA, 23219, 21<sup>ST</sup> FLOOR, PROCUREMENT OFFICE, TO THE ATTENTION OF MONIQUE ROBINSON (See Section VIII, C. Identification of Proposal Envelope.) This is NOT a mailing address. It is recommended that proposals be hand delivered.

In compliance with this Request For Proposals (RFP) and all conditions imposed in this RFP, the undersigned firm hereby offers and agrees to furnish all goods and services in accordance with the attached signed proposal or as mutually agreed upon by subsequent negotiation, and the undersigned firm hereby certifies that all information provided below and in any schedule attached hereto is true, correct, and complete.

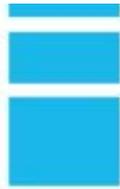
Name And Address Of Firm:

Northwest Evaluation Association (NWEA)  
121 NW Everett Street  
  
Portland, OR Zip Code: 97209  
eVA Vendor ID or DUNS #: E5209 / 60653577  
Fax Number: (503) 639-7873  
E-mail Address: proposals@nwea.org

Date: October 23, 2012  
By:   
(Signature In Ink)  
Name: Jeffrey P. Strickler  
(Please Print)  
Title: Executive Vice President & COO  
Telephone Number: (503) 624-1951

**PREPROPOSAL CONFERENCE:** An optional proposal conference will be held on Tuesday, October 9, 2012, at 10:00 am at 101 N. 14th Street, 25<sup>th</sup> Floor, Washington Conference Room (Reference Page 7, Section VII, herein). If special ADA accommodations are needed, please contact Marie Williams at 804 225-2040 by October 5, 2012.

**Note: This public body does not discriminate against faith-based organizations in accordance with the Code of Virginia, § 2.2-4343.1 or against a bidder or offeror because of race, religion, color, sex, national origin, age, disability, or any other basis prohibited by state law relating to discrimination in employment.**



NWEA™

Northwest Evaluation Association

*Partnering to help all kids learn®*

(This page left intentionally blank)



# COMMONWEALTH of VIRGINIA

## DEPARTMENT OF EDUCATION

P.O. BOX 2120  
RICHMOND, VA 23218-2120

October 17, 2012

### ADDENDUM NO.1 TO ALL OFFERORS

Reference – Request for Proposal: RFP #DOE-SGA-2012-15  
Commodity: 92420  
Dated: September 25, 2012  
For Delivery To: Department of Education  
Sealed Proposals Due: 2:00 PM October 29, 2012  
Pre-proposal Conference: October 9, 2012

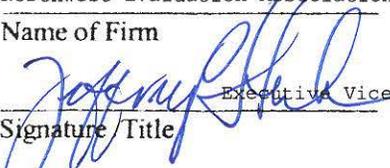
The above is hereby changed to read:

1. Reference Page 1, Using Agencies and Page 12, ADDITIONAL USERS: Add “, and consortia of Virginia Public School Divisions.”
2. Reference Page 5, B. SPECIFIC PROPOSAL INSTRUCTIONS: Add “6. Provide seven copies of each proposed assessment (for the evaluation panel and the procurement file). For assessments that are still under development, sample items may be submitted and should accompanied by additional information about the total number of items being developed for the proposed test.”
3. Reference Page 16, XI. PRICING SCHEDULE: Replace the title of column D with “Fixed Price per Year per Division.” Replace the title of Column E with “Unit Price (per test or per student).” Replace the title of column F with “Total Proposed Price for Scenario (Column D + (5,000 \* Column E)). Replace the first sentence with the following: “For each assessment offered, using the scenario of a minimum of 5,000 units (tests or students, as stated in the proposal) STATEWIDE per year, the required information and pricing per year must be provided using the table layout below. The unit price entered in column E may be in addition to a fixed price entered in column D or instead of a fixed price entered in column D. If only a unit price is offered, “0” should be entered in column D. If only a fixed price is offered, “0” should be entered in column E. DOE will evaluate price based on the amount to be entered in column F (final fixed price + (the unit price times 5,000)). A school division who chooses to use the contract for a given year would actually pay the final fixed price + the final unit price times the number of tests actually administered.”

Note: A signed acknowledgment of this addendum must be received at the location indicated on the RFP either prior to the proposal due date and hour or attached to your proposal. Signature on this addendum does not substitute for your signature on the original proposal document. The original proposal document must be signed.

Northwest Evaluation Association (NWEA)

Name of Firm

  
Executive Vice President & COO

Signature / Title

October 23, 2012

Date

Sincerely,



Marie G. Williams, VCO

Director of Accounting/Acting Director of Procurement



(This page left intentionally blank)

## Pricing Schedule

A. Assessment Description Attachment #	B. Name of Assessment	C. Content Area	D. Fixed Price per Year per Division	E. Unit Price (per test or student)	F. Total Proposed Price for Scenario (Column D + (5,000 * Column E))
Computer-adaptive assessment for grades 2-12 (Attachment 3)	MAP for Mathematics, Reading, and Language Usage Assessments	Mathematics, Reading, and Language Usage	\$4,360.00 <sup>1</sup>	\$10.80	\$58,360.00 <sup>2</sup>
Computer-adaptive assessment for grades 3-10 (Attachment 4)	MAP for Science Assessments <sup>3</sup>	Science	\$0.00	\$2.50	\$12,500.00 <sup>2</sup>
Computer-adaptive assessment for grades K-2 (Attachment 5)	MAP for Primary Grades Assessments	Mathematics and Reading	\$4,360.00 <sup>1</sup>	\$10.80	\$58,360.00 <sup>2</sup>

<sup>1</sup> Fixed Price includes \$4,360 for one MAP Administration Professional Development session (\$1,360 DMBE provided materials, \$3,000 MAP Administration Workshop).

<sup>2</sup> Proctor Training is embedded in the Unit Price (\$304 DMBE provided training).

<sup>3</sup> MAP for Science assessments must be purchased with MAP for Mathematics, Reading, and Language Usage assessments. These assessments cannot be purchased separately.



NWEA™

Northwest Evaluation Association

*Partnering to help all kids learn®*

(This page left intentionally blank)

# Proposal Narrative

## Overview

The measurement of student growth is an important aspect of evaluation in terms of understanding what students know and, perhaps more importantly, how they can improve. Growth data may be used to identify problem areas early and when used properly may help educators more effectively support student learning. Interim growth measurement provides educators a chance to check their own abilities as well, to see whether or not their practices are positively affecting their students.

The Virginia Department of Education recognizes the value and utility of these assessments and, with the release of an RFP for student growth assessments, demonstrates a dedication to helping students and teachers understand the how they may improve. Additionally, the assessments selected through this RFP will be a serve as a pathway to move divisions towards growth-focused measures.

Northwest Evaluation Association™ (NWEA™) has worked with Virginia educators and LEAs to measure growth for over eleven years. We currently provide growth assessments in thirteen public school divisions and administered 262,828 tests within the Commonwealth during the 2011-2012 school year. Throughout this work, we have helped educators in the Commonwealth understand how students are progressing towards success on the Virginia Standards of Learning (Virginia SOL) and assess student learning and growth across the school year in support of student learning. We commend the Department for releasing this RFP in acknowledgment of this important aspect of assessment and sincerely hope we can continue to help educators and students across the Commonwealth.

### *Proposal for the Virginia Department of Education*

In consideration of your goals to provide divisions with a stable, growth-oriented teacher accountability tool, as well as to empower educators with effective growth assessments, NWEA would like to recommend the Department consider our Measures of Academic Progress® (MAP®) program for approved use in the Commonwealth of Virginia. MAP is a computer-adaptive assessment that leverages a well-established scale to measure growth in an efficient and reliable way. The MAP program combines four powerful features – adaptive technology, assessment content, a stable vertical measurement scale, and educator resources – to provide Virginia educators with a reliable growth measure for all students and a foundation for educator evaluations.

MAP is a fully adaptive assessment, which allows it to provide a more detailed and reliable measure of student ability than any standalone or fixed-form test. MAP is not a patchwork of fixed form tests – it adapts at the item level to provide a truly adaptive experience for each student and a richer measurement of student ability.

In addition to being adaptive, MAP assessments are also capable of measuring student ability in standards above and below grade level. This is a tremendous benefit for both educators and students, especially with regard to growth. A fourth grade student who is reading at the second

grade level taking a traditional interim assessment would simply fail. With MAP assessments, the student's teacher will see not only that the student is performing at the second grade level, but also be able to identify areas the student needs additional help in order to perform on-grade or better.

MAP for Mathematics, Reading, Language Usage, and Science assessments are currently aligned to the Virginia SOL and will be ready to implement upon approval. Aligned assessments for mathematics, reading, and language usage are available for use in grades 2-12. Aligned assessments for science are available in grades 3-10. Mathematics and reading assessments are also available in kindergarten through second grade with MAP for Primary Grades; however, these tests are currently aligned to national standards.<sup>1</sup> We are in the process of aligning these tests to the Virginia SOL with aligned assessments for K-2 available in spring 2013.

### *Stable Testing Platforms*

NWEA's stable technology platform offers divisions an integrated system for delivering assessments, managing data, and accessing reports – all available from a web browser. This web-based system includes new capabilities, while maintaining current functionality and the value of the data NWEA's quality assessments provide. We are aware not all Virginia divisions have the bandwidth to handle the volume of data required by a web-based solution. As a result, we will be providing a client-server solution to those divisions until internet capacity is reached.

### *Teacher and Principal Evaluation*

NWEA recognizes that a purpose of this RFP is to identify assessments that will produce data to be used in teacher and principal evaluations each year. We acknowledge that there is a place for student assessment data as a factor in the evaluation of teachers or principals. In defining the appropriate use of these data for teacher or principal evaluation, we believe that schools should embrace existing research and known best practices in human resource management, and engage in collaborative efforts involving teachers, administrators, and governing bodies to strengthen teacher evaluation systems.

With these strong systems in place, NWEA supports the use of NWEA testing data as one factor among many to be considered by well-trained school leaders and other supervisors in the implementation of fair and rigorous expectations for teachers and principals. NWEA urges administrators to triangulate NWEA testing data with other sources of information for evaluation purposes, and recommends that student assessment data be used to inform a teacher's evaluation, and should not be the majority or controlling factor in the final evaluation of the performance of a teacher or principal. Measures of student achievement and growth, such as MAP and MAP for

---

<sup>1</sup> MAP for Primary Grades' (K-2) mathematics test has been aligned to the National Council of Teachers of Mathematics (NCTM) standards. MAP for Primary Grades' (K-2) reading test has been aligned to the National Council of Teachers of English (NCTE) standards. This alignment takes into account the International Reading Association Standards, National Reading Panel Report, and National Research Council Report "Preventing Reading Difficulties in Young Children."

Primary Grades assessments, are not by themselves indicative of a teacher's effectiveness; in other words, a measure of average raw student growth by itself does not constitute evidence of a teacher's effectiveness. For that inference to be made, student growth data must be translated to a rating of teacher effectiveness, typically using some form of value-added analysis. Because the analytic approaches to this process differ, the choices made around the methodology or procedures used can have a substantive effect on the validity of inferences made about a teacher.

Further, because progress measures require multiple tests, the conditions in which these tests are administered can also impact inferences about a teacher's effectiveness. For example, the MAP test is supposed to be untimed, so if the test was administered in a compressed time period (such as during the half-hour before lunch), the validity of the results of the assessment could be seriously compromised. This could also invalidate inferences made about a teacher's effectiveness.

The point is that no student achievement test can be validated as a test of a teacher's effectiveness per se. Instead, a test that provides an accurate measure of student achievement – such as MAP and MAP for Primary Grades assessments – can provide data that may be used to reasonably estimate a teacher's contribution to learning. However, in order for this to occur, the test must be properly administered, and consideration must be given to the ways in which student test scores are translated to an evaluation of teacher performance.

NWEA offers general guidance that school leaders should consider before using MAP or MAP for Primary Grades assessments as an evaluation system. For example, NWEA does not recommend the use of MAP or MAP for Primary Grades assessment data in the evaluations of teachers who teach non-tested subjects, such as art or music. Because MAP and MAP for Primary Grades assessments are not aligned to the core professional duties of teachers in these fields, it would be invalid to evaluate a teacher's job performance on an assessment that is not aligned to their teaching assignment. Further, for all teachers, it is necessary to ensure that there is alignment between the content of the relevant subject areas taught in the classroom and the content that is assessed in those same subject areas on the MAP or MAP for Primary Grades assessments, so that all evaluations of performance are fair reflections of the work of the teacher.

There are a number of complex issues to consider surrounding the use of student tests as measures of teacher effectiveness. Because of this, where MAP and/or MAP for Primary Grades assessments are approved for use in an evaluation system, NWEA will offer technical guidance and support to ensure appropriate use of NWEA testing data. NWEA will also work collectively with divisions to devise a state-wide recommendation for how student testing data could be used in the evaluation of teacher and principal performance. This will help make certain decisions made about teacher performance based on student testing data are carefully considered and technically sound. This can include our recommended use of a value-added model provided by a third party. NWEA has worked with a number of value added partners and would gladly share our experience and advice in using these models in Virginia.

### Conclusion

In our proposal, you will find greater detail about our proposed MAP assessment and highlights of its effectiveness in delivering reliable, valid measurement of growth. We believe that these assessments will meet and exceed your expectations. We confidently stand by the quality of data our computer-adaptive assessments produce and support the Commonwealth's use of this data in teacher and principal evaluation. We look forward to continuing our eleven-year history of working with educators in Virginia by gaining your approval of our assessments.

### a. Organizational Structure

NWEA started in 1974 as a joint venture of representatives from Pacific Northwest school districts, including Portland and Seattle. NWEA initially focused on development of its Rasch Unit (RIT) scale and accompanying items. In 1977, NWEA was incorporated as a non-profit corporation, and in 1986 became the first organization to offer computer-delivered testing to K-12 educational agencies. Since delivering its first computerized adaptive test in 1986, NWEA has continued to improve its assessments, research activities, and professional development services.

Since becoming an incorporated non-profit organization in 1977, NWEA has grown to employ more than 500 highly qualified and dedicated staff members. While the organization's headquarters are in Portland, Oregon, NWEA has offices in New York City, New York and Pottsboro, Texas. Additionally, the organization supports a number of staff working from home offices located across the country.

The organization is divided into eight departments, which are managed by vice presidents that report directly to either NWEA's CEO, Matt Chapman, or Executive Vice President and COO, Jeffrey Strickler. See Appendix A to view the organizational structure of NWEA.

NWEA uses the term *partners* rather than *clients* because partnership more accurately describes its relationships with the people who use its products and services.

#### *Development of the MAP Assessments*

Dedicated to the belief that assessments should make a difference for students, NWEA developed MAP, a computer-adaptive assessment that provides educators with information they can use to improve teaching and learning. MAP assessments being made available for Virginia educators include:

- MAP for Mathematics, Reading, and Language Usage Assessments for Grades 2-12
- MAP for Science Assessments for Grades 3-10
- MAP for Mathematics and Reading Assessments for Grades K-2

Purchase of one of the MAP assessments listed above includes: online reports and data files, technical support, implementation support, ongoing account management, and online resources.

### *Technical Support*

NWEA's Technical Support Team is available by toll-free Support line or email, Monday through Friday, 7:00 a.m. to 5:00 p.m. Eastern Time (ET), excluding federally-recognized holidays observed by NWEA. The Technical Support staff is well-versed in the implementation of NWEA's assessment system and can provide assistance with generating Roster Files, configuration of the MAP system components, accessing online reports, and questions that may arise in the use of the system.

NWEA's call routing and escalation protocols efficiently route service requests to the appropriate personnel. NWEA staff will respond with accurate, timely, courteous, and consistent service. To optimize responsiveness and maintain consistently high customer satisfaction, NWEA uses established Information Technology Infrastructure Library-based (ITIL) escalation protocols.

During a division's first implementation testing season, NWEA will assign a dedicated Implementation Support Specialist to proactively guide designated LEA staff through the steps for preparing for a test season and retrieving online reports. The Implementation Support Specialist maintains continuous contact before, during, and at the conclusion of the first testing season and is available to answer questions throughout the testing season.

After a partner has completed the first MAP testing season, ongoing account management duties are transitioned from the Implementation Support Specialist to the Account Manager. Whereas NWEA's implementation support services focus on helping the partner work through the logistics of technical setup and test administration, the Account Manager schedules periodic check-in meetings with the partner to answer questions, follow up on any open issues, schedule additional professional development as needed, and collaborate with LEA staff on their plans for the next year of testing.

As a complement to standard technical support services, NWEA also offers optional Technical Consulting services which can be purchased separately by partners to assist with accounts that have a customized or complex technical configuration. This team of highly trained technology professionals works closely with the Account Manager and Implementation Support Specialist to resolve any issues involved with the technical setup of the MAP software at partner sites. The Account Manager, Implementation Support Specialist, and Technical Support staff may refer a partner to this team for customized assistance. NWEA's Technical Consultants may be on the ground with a partner, provide phone support and remote desktop support, and/or coordinate the acquisition of other resources to support a successful testing experience from a technology point of view.

## **b. Maintaining Confidentiality of Student Data**

All student information is securely transferred over encrypted channels. For the purposes of transferring assessment information at the completion of each administration period, data are made available to designated division personnel as a .CSV (comma-separated values) file or one or more spreadsheets accessible from our secure FTP server. Confidentiality of all assessment

information, including responses to assessment items by students as well as assessment reports for school officials, is ensured by using high-grade encryption whenever information is transferred over a network, including data transferred in the course of the administration of an assessment and while reports are being accessed by school officials via the web-based report software. All information collected by NWEA is stored and backed-up on a secure network audited for HIPAA and COPPA compliance. Access to student information is restricted on a need-to-know basis by NWEA employees. NWEA's Information Handling Standards are included as Appendix B.

### **c. NWEA's Training Plan**

#### *On-Site Technical Training and Professional Development*

NWEA has a strong commitment to professional development for central office staff, teachers, and administrators. As part of every program, we offer a fully developed set of learning solutions to fit LEA schedules and needs.

NWEA's MAP Administration workshop, delivered by our facilitators, is the key component of our training plan. MAP Administration workshops are hosted by the division and are facilitated within one day. The division is responsible for providing facilities and equipment (e.g., computers with Internet connection and projectors) for the delivery of the MAP Administration workshop. A MAP Administration workshop can accommodate up to forty participants per session.

Division staff will be required to assist NWEA in coordinating the onsite MAP Administration workshop. Coordination typically includes scheduling events, communication with workshop participants, meal planning, distributing pre-workshop materials, and tracking participant attendance.

During MAP Administration workshops, attendees learn how to implement and administer MAP tests and gain an understanding of how to use adaptive testing to measure student progress and identify instructional needs. This workshop occurs prior to the first test administration. See Appendix C for a comprehensive overview of NWEA's MAP Administration workshop.

#### *Enrichment for Virginia Educators*

##### **User Experience Meetings**

NWEA will provide User Experience meetings to give Virginia educators a better understanding of MAP and MAP for Primary Grades assessments. User Experience meetings are one-day events. Breakfast, lunch, and snacks, plus all printed materials are provided by NWEA. Within User Experience meetings, a number of breakout sessions are held. An example of User Experience sessions follows:

- **Growth Reports and Goal Setting:** This presentation centers on the variety of reports that are available at the teacher, school, and/or division levels for analyzing growth patterns and setting goals for the school and division using the growth patterns found, as well as the growth calculator.

- Reports Basics: An overview of the most frequently used NWEA reports. Focused discussion on defining reporting terms and describing various data points.
- Dynamic Reporting Suite: The Dynamic Reporting Suite gives teachers and administrators data on which instructional decisions can be based. This session introduces the tools and reports available in the Dynamic Reporting Suite.
- DesCartes: a Continuum of Learning<sup>®</sup> and Instructional Resources: These two tools can help division staff to make the connection from data to powerful instruction for all students. With Instructional Resources and DesCartes, teachers can identify critical gaps in their students' understanding and develop powerful lessons targeting students' readiness to learn.
- MAP for Primary Grades: These assessments provide teachers with an efficient way to assess achievement levels of early learners so they can spend more time teaching and less time administering individual diagnostic tests. This session focuses on how teachers can use MAP for Primary Grades to identify the needs of all primary grades students from struggling to advanced learners.
- The 2011 Norming Study: This session will present the newest norming study released by NWEA. It will review the types of normed information available to NWEA's partners and how it can best be used by school level personnel. This will be a "how to use this document" version and will not delve into the technical aspects of how the norms were completed/calculated, etc.

### **Consortium Conferences and Regional Research Presentations**

NWEA's Consortium Conferences will be three hours in length. Lunch and all printed materials will be included by NWEA. LEAs will be able to attend free of charge and learn about NWEA's MAP assessments. Focus will be placed on student growth measures, teacher evaluation, and specific Virginia initiatives relative to the audience. These conferences are facilitated by NWEA staff knowledgeable about current Virginia issues and psychometric best practices.

Regional Research Presentations will be three hours in length. Lunch and all printed materials will be provided by NWEA. LEAs will be able to attend free of charge and listen to a member of NWEA's research team about various topics relative to Virginia LEA's Growth model. NWEA's experience research staff will facilitate these presentations.

A tentative schedule of these conferences is included below:

*Proposal Narrative, c. NWEA's Training Plan*

Approximate Date	Regional Centers/Consortium	General Location (For Caterer)
January 2013	North Tier Consortium	Fredericksburg, VA
March 2013	Shenandoah Valley Consortium	Charlottesville, VA
May 2013	Four Rivers Consortium	West Point, VA
October 2013	Central Virginia Consortium	Richmond, VA
November 2013	WHRO/CII	Williamsburg, VA
May 2014	Blue Ridge East Consortium	Roanoke, VA
January 2014	Southside Consortium	Prince Edward, VA
March 2014	Blue Ridge West Consortium	Radford, VA
February 2013	Regional Research Presentation	Roanoke, VA
September 2013	Regional Research Presentation	Richmond, VA
February 2012	Regional Research Presentation	Alexandria, VA

***Proctor Training***

As part of the implementation of MAP assessments, NWEA provides supplemental proctor training. This training is delivered via a live, one hour webinar and is designed for division staff members that have completed the MAP Essentials online trainings, and need additional guidance on how to proctor MAP assessments. The Proctor training is facilitated by a member of the NWEA Implementation Support Staff. Participants will need access to a computer with internet access and a telephone to participate in the webinar. If the session will be viewed by multiple participants in one location, a speaker phone must be provided by the division or school hosting the training session. Prior to attending this webinar, participants should: 1) View the online MAP Essentials training sessions, including "Proctor Quick Start" and "Proctor Tools"; and 2) Review the "Proctor Tips and Troubleshooting" document.

***Building Local Expertise and Certified Facilitator Programs***

NWEA has successfully supported partners who have chosen to utilize the "Building Local Expertise" format by recommending that an instructional team made up of three representatives from each school in the division participate in the on-site workshops. Following the workshops, the instructional team takes responsibility for providing the workshop content, materials, tools, and resources to their peers by conducting teach-backs for their home school location. Our standard offering MAP Foundations Series workshops can accommodate up to forty participants per session.

In addition to the Building Local Expertise format of the MAP Foundation Series workshops, NWEA offers a Certified Facilitator Program in which local educators become certified to facilitate the Foundation Series workshops for their home divisions or for other divisions using the MAP system. Additional information and pricing for the Certified Facilitator Program is available upon request.

### *Optional Workshops*

NWEA's Professional Development team also provides a variety of other optional workshops designed to help teachers, administrators, and division staff members master the art of data-driven instruction. Our series of on-site workshops, known as the MAP Foundations Series, are available to NWEA's partners. These workshops focus on successfully implementing MAP assessments, and then utilizing the resulting MAP data to improve student achievement. The NWEA facilitators target instruction to meet the specific needs of workshop participants. These workshops can be delivered in a "turn-key" train-the-trainer model where additional resources and tools are provided to workshop participants, enabling them to take the content from the on-site workshops and conduct teach-backs for their home school locations.

NWEA's MAP Administration workshop (included in NWEA's Pricing Schedule) is typically the first workshop of the MAP Foundation Series. Other optional MAP Foundations Series one-day workshops include:

- **Stepping Stones to Using Data –** Analyze and learn to interpret data, and engage with other faculty to create an environment responsive to the needs of students. Data at the student, classroom, school, and division levels are explored and interpreted for practical application in this workshop. Ideally, this workshop should occur between the first and second test administrations so that partners have their own assessment data which they can refer to and apply.
- **Climbing the Data Ladder –** Apply the information from reports to instructional practices and use test results to differentiate instruction, form flexible groups, and develop strategies to meet each student's needs. Participants learn more about using instructional resources such as DesCartes: A Continuum of Learning to differentiate instruction and improve learning.
- **Growth and Goals –** Learn how to work with students to set growth goals and communicate effectively with all stakeholders. School and division administrators will learn to analyze data over time for effective program management. Participants in this workshop should have a minimum of two test administrations' worth of data to be able to successfully reference and analyze their own longitudinal assessment results.

### *Online Technical Training*

MAP system users have access to online product training designed to accelerate their proficiency with the assessment. These interactive, self-paced training modules cover the following topics: introduction to the MAP system, setting up the system, how to proctor assessments, and understanding reports.

Additional online resources include SPARK (<http://community.nwea.org/>), NWEA's online forum, which provides a means for educators to share best practices and success stories, learn from their peers in education and NWEA staff, and obtain resources that support meaningful instruction and the promotion of kids' growth and learning.

#### **d. Using MAP Assessments as Measures of Growth**

MAP assessments have been selected worldwide as a key growth indicator by numerous state and local education agencies. Educators in these organizations have chosen MAP to measure student growth for a number of different purposes.

We have included here a sample of our excellent performance in statewide evaluations of our instruments.

##### *Review and Approval of MAP for Mathematics, Reading, and Language Usage to Measure Growth*

In August 2011, MAP for Mathematics, Reading, and Language Usage assessments were reviewed and approved by the New York State Education Department's list of approved growth assessments of the New York State Common Core Standards for use in teacher and principal evaluations. The list of approved assessments including MAP tests can be viewed on the following website: <http://usny.nysed.gov/rttt/teachers-leaders/assessments/home.html#assess>.

In April 2012, MAP for Mathematics, Reading, and Language Usage assessments were reviewed and approved by the Ohio Department of Education's list of approved assessments that can be used for a portion of teachers' and principals' evaluations. To be approved, NWEA submitted evidence that MAP assessments met the fundamental requirements for measuring student growth as identified by the Ohio Department of Education:

1. Be highly correlated with curricular objectives
2. Have enough "stretch" to measure the growth of both low-and high-achieving students
3. Meet appropriate standards of test reliability.

##### *Review and Approval of MAP for Science to Measure Growth*

In August 2011, MAP for Science assessments were reviewed and approved by the New York State Education Department's list of approved growth assessments of the New York State Common Core Standards for use in teacher and principal evaluations. The list of approved assessments including MAP tests can be viewed on the following website: <http://usny.nysed.gov/rttt/teachers-leaders/assessments/home.html#assess>.

In April 2012, MAP for Science assessments were reviewed and approved by the Ohio Department of Education's list of approved assessments that can be used for a portion of teachers' and principals' evaluations. To be approved, NWEA submitted evidence that MAP assessments met the fundamental requirements for measuring student growth as identified by the Ohio Department of Education:

1. Be highly correlated with curricular objectives
2. Have enough "stretch" to measure the growth of both low-and high-achieving students
3. Meet appropriate standards of test reliability.

### **Review and Approval of MAP for Primary Grades to Measure Growth**

In January 2012, MAP for Primary Grades assessments were reviewed and approved by the New York State Education Department's list of approved growth assessments of the New York State Common Core Standards for use in teacher and principal evaluations. The list of approved assessments including MAP tests can be viewed on the following website:

<http://usny.nysed.gov/rttt/teachers-leaders/assessments/home.html#assess>.

### **e. Demonstrated Experience**

More than 6,000 education agencies worldwide have selected NWEA's solutions because of its high quality, utility, and accuracy. NWEA works closely with many of these stakeholders to target their assessment programs to best address the needs of current and future learners. The close personal service provided to each of these partners by the entire organization is an integral element of these programs. NWEA's strong assessments, partner support services, and professional development are evidenced by the longevity in the contracts and the level of satisfaction these partners relay to NWEA and the assessment community.

Because of the large number of partners (particularly those district-level partners selecting MAP assessments over the past three years) currently using our assessments, a brief listing, including descriptions of NWEA's quality partner experiences in programs comparable to the Virginia Student Growth Assessments, is provided below.

#### **Chicago Public Schools (2008 – Present)**

NWEA started a partnership with the Chicago Public Schools (CPS) Autonomously Managed Performance Schools (AMPS) program during 2008, working with CPS AMPS to provide computer adaptive assessments, training, professional development, support (implementation and ongoing technical), and program management. The initial pilot began with training seventy staff in fifteen schools to assess approximately 6,000 students.

In 2009, NWEA was awarded a contract with CPS to allow NWEA to expand throughout AMPS. During the 2010-2011 school year, AMPS had expanded testing to more than 50,000 students in more than 100 schools. During the 2011-2012 school year, CPS tested more than 63,000 students using NWEA's client-server and Web-based assessment solutions.

NWEA's Research Department supports CPS through custom data and research services. This work includes consultative services on how to work with and apply the data. The success of this partnership is largely due to frequent and ongoing communication between NWEA program management and AMPS leadership and staff around data analysis, program evaluation, and general partnership needs.

In February 2012, NWEA was awarded the contract for Adaptive Growth Assessments. As part of the contract, all CPS students in grades K-8 use NWEA assessments as of fall 2012. The district plans to test more than 220,000 students in approximately 350 schools using NWEA's web-based assessments in 2012-2013.

### **Milwaukee Public Schools (2006 – Present)**

In 2006, Milwaukee Public Schools became an NWEA partner. Of the district's 87,000 students, 62,000 are licensed to test using NWEA's assessments. During the 2011-2012 school year, the district administered more than 476,500 assessments to its students.

Milwaukee Public Schools uses NWEA's assessments as a screener for its students. NWEA's assessments have afforded the district's teachers, students, and parents the opportunity to gain valuable and immediate information regarding student progress three times per year. This allows for immediate intervention when necessary and celebration of expected and unexpected gain.

### **Plano Independent School District (2000 – Present)**

In 2000, a Plano Independent School District (ISD) and NWEA partnership began. Of Plano ISDs more than 55,000 students, 37,936 are licensed to test with NWEA's assessments.

NWEA assessment data are used for a variety of purposes including instructional focus, accountability, teacher incentive programs, and to monitor the academic growth of Plano ISDs students. Plano ISD also developed a local based reporting system used by its teachers and administrators that effectively tracks NWEA assessment data so that each and every student can be tracked for academic growth, college readiness, and interventions. Plano ISD also created early warning systems using NWEA assessment data to track the success of its students on the state summative test.

### **Utah State Office of Education (2009 – Present)**

A pilot in spring 2009 administered customized MAP tests designed for the purpose of providing both summative and interim data in a single assessment (the Utah Local Adaptive Assessment (ULAA)) in an effort to reduce the number of assessments administered to Utah students. In the second and third years of the pilot, additional schools and LEAs participated to increase the sample size.

The Blended Assessment program is an adaptive test combining both interim and summative assessments into a single test administration. The first segment presents items within the grade level and produces the student's proficiency level. The second segment provides additional detail about student achievement and growth regardless of the student's grade level.

Through the USOE, NWEA's Blended Assessment was submitted for Federal Peer Review and was conditionally approved for each in 2010, 2011, and 2012. The USOE is currently awaiting results from the Peer Review Committee in anticipation of receiving final approval. Ultimately, once the USOE received unconditional approval, all Utah public LEAs would have the option of using the ULAA as an alternative to the CRT.

## **f. NWEA's References**

NWEA submits the following references:

1. Dr. Penny Blumenthal  
Director of Research and Planning  
Henrico County School District  
3820 Nine Mile R  
Richmond, VA 23223  
Tel: (804) 652-3600  
[mdblumen@henrico.k12.va.us](mailto:mdblumen@henrico.k12.va.us)
2. H. Alan Seibert, Ed. D.  
Division Superintendent  
Salem City Schools  
510 South College Avenue  
Salem, VA 24153  
Tel: (540) 389-0130  
[aseibert@salem.k12.va.us](mailto:aseibert@salem.k12.va.us)
3. Jamie Mathieson  
Director of Testing and Accountability  
Fluvanna Public Schools  
14455 James Madison Highway  
Palmyra, VA 22962  
Tel: (434) 589-8208  
[jmathieson@apps.fluco.org](mailto:jmathieson@apps.fluco.org)  
See Appendix D for a letter of reference.



(This page left intentionally blank)

# Attachment 1 – Small Business Subcontracting Plan

## Definitions

**Small Business:** "Small business " means an independently owned and operated business which, together with affiliates, has 250 or fewer employees, or average annual gross receipts of \$10 million or less averaged over the previous three years. Note: This shall not exclude DMBE-certified women- and minority-owned businesses when they have received DMBE small business certification.

**Women-Owned Business:** Women-owned business means a business concern that is at least 51% owned by one or more women who are citizens of the United States or non-citizens who are in full compliance with United States immigration law, or in the case of a corporation, partnership or limited liability company or other entity, at least 51% of the equity ownership interest is owned by one or more women who are citizens of the United States or noncitizens who are in full compliance with United States immigration law, and both the management and daily business operations are controlled by one or more women who are citizens of the United States or non-citizens who are in full compliance with the United States immigration law.

**Minority-Owned Business:** Minority-owned business means a business concern that is at least 51% owned by one or more minority individuals or in the case of a corporation, partnership or limited liability company or other entity, at least 51% of the equity ownership interest in the corporation, partnership, or limited liability company or other entity is owned by one or more minority individuals and both the management and daily business operations are controlled by one or more minority individuals.

**All small businesses must be certified by the Commonwealth of Virginia, Department of Minority Business Enterprise (DMBE) by the due date of the solicitation to participate in the SWAM program. Certification applications are available through DMBE online at [www.dmbv.org](http://www.dmbv.org) (Customer Service).**

Offeror Name: Northwest Evaluation Association (NWEA)

Preparer Name: Michael Hurst Date: 10/22/2012

## Instructions

- A. If you are certified by the Department of Minority Business Enterprise (DMBE) as a small business, complete only Section A of this form. This shall not exclude DMBE-certified women-owned and minority-owned businesses when they have received DMBE small business certification.
- B. If you are not a DMBE-certified small business, complete Section B of this form. For the offeror to receive credit for the small business subcontracting plan evaluation criteria, the offeror shall identify the portions of the contract that will be subcontracted to DMBE-certified small business in this section. Points will be assigned based on each offeror's proposed subcontracting expenditures with DMBE certified small businesses for the initial contract period as indicated in Section B in relation to the offeror's total price.

## Section A – *Not Applicable (NA)*

If your firm is certified by the Department of Minority Business Enterprise (DMBE), are you certified as a (check only one below):

Small Business

Small and Women-owned Business

Small and Minority-owned Business

Certification number: \_\_\_\_\_ Certification Date: \_\_\_\_\_

**Section B**

Populate the table below to show your firm's plans for utilization of DMBE-certified small businesses in the performance of this contract. This shall not exclude DMBE-certified women-owned and minority-owned businesses when they have received the DMBE small business certification. Include plans to utilize small businesses as part of joint ventures, partnerships, subcontractors, suppliers, etc.

**C. Plans for Utilization of DMBE-Certified Small Businesses for this Procurement**

Small Business Name & Address	Status if Small Business is also: Women (W), Minority (M)	Contact Person, Telephone & Email	Type of Goods and/or Services	Planned Involvement During Initial Period of the Contract	Planned Contract Dollars During Initial Period of the Contract
Bayside Printing, LLC 2681 Waldrop Church Rd. Loisa, VA 23093 Cert. #: 257	W	Marcia Wills Ph: (540) 223-8525 <a href="mailto:VAformslady@aol.com">VAformslady@aol.com</a>	Materials for State-wide users experience meetings	Provide printed materials for eleven scheduled trainings	\$8,400.00
Conference Solutions, LLC. PO Box 7276 Fairfax Station, VA 22039 SWAM Cert. #: 646954	W	Jennifer Eichel Ph: (703) 493-9287 <a href="mailto:jeichel@customconference.com">jeichel@customconference.com</a>	Training coordination services for user experience meetings	Notify attendees, reserve facilities, coordinate with caterers, etc.	\$25,000.00
Holder Enterprises, Inc. 457 Ride St. NW, Suite 1 Washington, DC 20001 Cert. #: DB20090144-2012	M	Selwyn E. Holder Ph: (202) 347-9545	Catering for statewide user experience meetings	Provide meals for eleven scheduled trainings	\$420.00
Customer Value Partners 3701 Pender Dr, Suite 200 Fairfax, VA 22030 Cert. #: 698546	M	Anirudh Kulkarni Ph: (703) 345-9118	Statewide Program Management Services	Manage DMBEs, complete monthly DMBE reports, develop performance indicator reports	\$20,000.00
Indrasoft, Inc 11150 Sunset Hills Rd., Suite 120 Reston, VA 20190 Cert. #: 693127	W	Neeraja Lingam Ph: (703) 435-9052	Division-level proctor training	Provide proctor trainings for divisions as part of the implementation process	\$304.00

Small Business Name & Address	Status if Small Business is also: Women (W), Minority (M)	Contact Person, Telephone & Email	Type of Goods and/or Services	Planned Involvement During Initial Period of the Contract	Planned Contract Dollars During Initial Period of the Contract
Bayside Printing, LLC 2681 Waldrop Church Rd. Loisa, VA 23093 Cert. #: 257	W	Marcia Wills Ph: (540) 223-8525 <a href="mailto:VAformslady@aol.com">VAformslady@aol.com</a>	Printed materials for MAP Administration	Print and ship materials for MAP Administration	\$1,360.00
Totals \$					\$55,484.00

NWEA is dedicated to supporting DMBE-certified businesses when it provides student growth assessments in Virginia. In the preparation of this proposal, NWEA identified businesses to help support NWEA's work providing student growth assessments. Given the general scope of this work however, we believe that the estimates presented in the subcontractor plan may vary in execution. Specifically, we anticipate that the amount of work (and in turn costs) associated with DMBE-certified businesses will vary based on the extent that NWEA's assessments are adopted by divisions in Virginia because much of total cost for assessments is based on a per unit pricing. Accordingly, the fixed amount we plan to dedicate to DMBE-certified businesses cannot be guaranteed in performance of a contract resulting from the RFP.

NWEA submits this plan on the good faith assumption that these businesses provide the high quality service NWEA expects from all its vendors. NWEA reserves the right without penalty from Virginia Department of Education to discontinue its contractual relationship with any DMBE-certified business if the business is unreliable or produces poor quality work, including cases in which the degree of unreliability or poor quality does not rise to the extreme of a contractual default.

NWEA acknowledges and accepts the obligation to record and report its use of DMBE-certified businesses to the Virginia Department of Education per the RFP.

Please note, that some of the work listed in NWEA's proposal relates to state-level work, like program management, that is not division specific. For example, NWEA intends to host a series of user group meetings across the state in 2013-2014. NWEA intends to use DMBE-certified businesses to support these efforts, but the cost of these user groups will not pass through to the divisions. We consider the support of such businesses related to providing services under the contract resulting from the RFP and intend to reflect our use of these businesses for state-level work as a part of NWEA's compliance with its Small Business Contracting Plan.



(This page left intentionally blank)

### State Corporation Commission Form

**Virginia State Corporation Commission (SCC) registration information. The offeror:**

is a corporation or other business entity with the following SCC identification number: F154319-0

**-OR-**

is not a corporation, limited liability company, limited partnership, registered limited liability partnership, or business trust **-OR-**

is an out-of-state business entity that does not regularly and continuously maintain as part of its ordinary and customary business any employees, agents, offices, facilities, or inventories in Virginia (not counting any employees or agents in Virginia who merely solicit orders that require acceptance outside Virginia before they become contracts, and not counting any incidental presence of the offeror in Virginia that is needed in order to assemble, maintain, and repair goods in accordance with the contracts by which such goods were sold and shipped into Virginia from offeror's out-of-state location) **-OR-**

is an out-of-state business entity that is including with this proposal an opinion of legal counsel which accurately and completely discloses the undersigned offeror's current contacts with Virginia and describes why those contacts do not constitute the transaction of business in Virginia within the meaning of § 13.1-757 or other similar provisions in Titles 13.1 or 50 of the Code of Virginia.

**\*\*NOTE\*\*** >> Check the following box if you have not completed any of the foregoing options but currently have pending before the SCC an application for authority to transact business in the Commonwealth of Virginia and wish to be considered for a waiver to allow you to submit the SCC identification number after the due date for proposals (the Commonwealth reserves the right to determine in its sole discretion whether to allow such waiver):



NWEA™

Northwest Evaluation Association

*Partnering to help all kids learn®*

(This page left intentionally blank)

**Offeror Name:** Northwest Evaluation Association (NWEA)

**Proposed Assessment Name:** MAP for Mathematics, Reading, and Language Usage Assessments

**Content Area(s) and Grade Level(s) Assessed:** Mathematics, Reading, and Language Usage for Grades 2-12

## Section 1: Overview of Tests

### Requirement: Section 1: Overview of Tests 1.1

#### 1.1

Describe the specific grade(s) and subject area(s) covered by each assessment and provide an overview of the content and skills measured. Include the types of test items used, the mode(s) of delivery, the availability of equivalent forms, including short forms or screeners (if available) and a test blueprint for each test being proposed.

MAP for Mathematics, Reading, and Language Usage assessments are computer adaptive tests aligned to the Virginia Standards of Learning (Virginia SOL) and contain selected-response items that are scored immediately. For mathematics and reading, there are two tests: a grade 2-5 test and a grade 6+ test. There is one Language Usage assessment for students in grades 2-12. There is a continuous scale between the two grades ranges, ensuring continuous measurement of student performance. The following test types are available for Mathematics, Reading, and Language Usage:

- Survey test – a twenty-item adaptive test that gives an overall RIT score (see Section 2.2.2 for a complete description of NWEA’s RIT scales). Survey tests typically take about thirty minutes to administer, and are most commonly used for course placement and intake testing.
- Survey with Goals test – adaptive tests designed to measure achievement of students as they grow through the grades. Survey with Goals tests typically consist of more than forty questions and provide an overall RIT score for the subject as well as sub-scores in each of the goal performance areas. The accuracy of these tests allows the results to be used to measure student growth as well as current achievement level. MAP Survey with Goals tests can be administered up to four times per year, and provide information about students’ achievements levels and their changes in performance over time.
- In mathematics NWEA also offers end-of-course tests in Algebra I, Algebra II, Geometry, Integrated Mathematics I & II, and Integrated Mathematics III. End-of-Course tests give an overall RIT score in a specific content area.

The content and skills covered by NWEA’s Virginia State-aligned MAP for Mathematics, Reading, and Language Usage assessments are demonstrated in Appendix E, which contains NWEA’s Goal Structure Charts. Appendix E also includes an explanation of how the structures of the Virginia SOL for each subject area were addressed in the development of the Virginia State-aligned MAP for Mathematics, Reading, and Language Usage assessments. See Appendix F for the MAP for Mathematics, Reading, and Language Usage Test Blueprint.

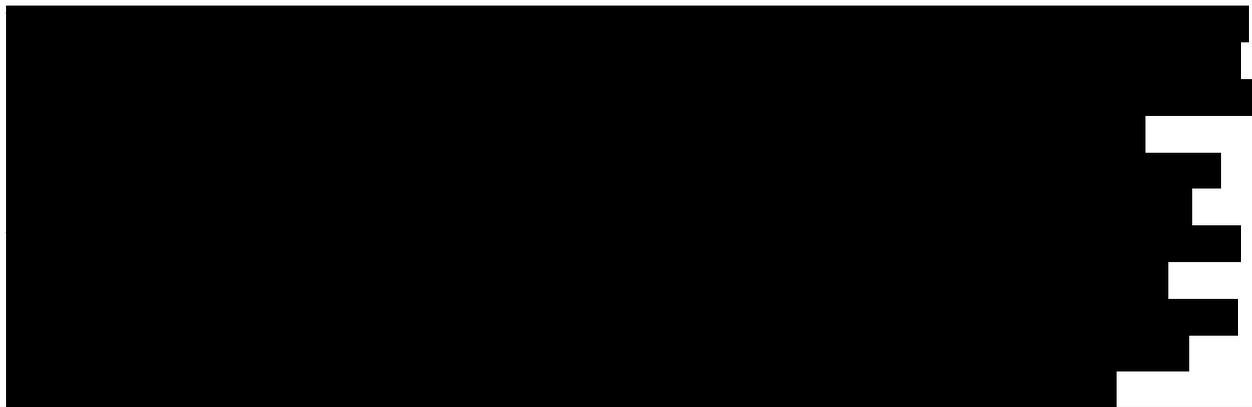
**Requirement: Section 1: Overview of Tests 1.2**

1.2

Provide evidence of alignment of test items to the Virginia Standards of Learning (SOL) for existing assessments. For assessments developed in response to the RFP, provide a plan for assuring the alignment of test items to the SOL.

The MAP assessments and associated reports are based upon test-specific goal structures created by NWEA Content Specialists. These Content Specialists are content area experts with five or more years of classroom teaching experience. They have an expert level of knowledge of writing and reviewing items for classroom, district, interim, and state assessments; have an expert knowledge of national trends and initiatives in assessment. Additionally, they have experience creating test structures aligned to state standards and selecting item pools for those assessments. In the alignment process, NWEA Content Specialists reviewed the Virginia SOL standards carefully and organized them into reporting frameworks (Goal Structures) for MAP tests. A goal structure was created by combining Virginia's standards and grade-level expectations across grade levels to create a two-tier framework consisting of goals and sub-goals. For example, the grades 2-5 MAP for Mathematics test has a "Number and Number Sense" goal that includes two sub-goals: "Whole Number Concepts" and "Fraction and Decimal Concepts." These goals and sub-goals have a direct relationship with the Virginia SOL.

Once the goal structures were finalized, NWEA Content Specialists aligned items from NWEA's item bank to each subgoal. This ensured the item pools for the Virginia assessments contain only topics within the Virginia SOL standards.



## Section 2: Technical Characteristics

### Requirement: Section 2: Technical Characteristics 2.1

2.1

Provide evidence of content, construct, concurrent, and predictive validity as appropriate. Include validity evidence that supports the use of scores from the proposed assessment in teacher evaluation, addressing specifically the validity of using assessment results to support inferences about effectiveness of teacher in producing growth in student performance (if available).

### **Requirement: Section 2: Technical Characteristics 2.1**

Many types of information can be used as validity evidence. This information ranges, for example, from the adequacy and coverage of a test's content, to its ability to yield scores predictive of a status in some area, to its ability to draw accurate inferences about a test taker's status with respect to a construct, to its ability to allow generalizations from test performance within a domain to like performance in the same domain. Some common forms of validity evidence can be offered to support broad validity arguments. Appendix H includes information about content, concurrent, predictive, and criterion-related validity. The vast preponderance of this evidence comes from the relationships of MAP test scores to state content-aligned accountability test scores. These come in several forms including:

- the test content;
- the concurrent performance of students on MAP tests with their performance on state tests given for accountability purposes;
- the predictive relationship between students' performance on MAP tests with their performance, two testing terms later, on state accountability tests; and
- the relationship between students' performance on MAP tests and their nominal status relative to criteria defined by their state's achievement standards.

Where student performance data are available from MAP tests aligned to Virginia SOL, those data are used to compute correlations coefficients to support validity arguments. In those cases where sufficient Virginia student performance data were not available for a particular type of validity argument, performance data from other states are presented as an example of how MAP state content-aligned test performance corresponds to performance on other (state accountability) tests.

MAP assessments are aimed at providing estimates of individual students' academic achievement that minimize measurement error, whether that error is directional in some form of bias or the result of unstable scales. NWEA warrants MAP tests to be valid for uses related to these characteristics. Such uses may entail, for example, making decisions about a student's achievement status, or about a student's change(s) in achievement status (growth), or understanding how a student's achievement status and growth compare to carefully constructed national norms. We have no conclusive or even systematically collected evidence about what antecedents contribute to or detract from a student's achievement estimates. We suspect that such relationships will be shown to be multivariate and highly multidimensional. Whatever the structure of these relationships, strong and convincing evidence about them will be fundamental to any argument that attributes a particular outcome to a particular antecedent or set of antecedents. Due to the absence of such basic evidence, we have not constructed validity evidence that supports the use of scores from the proposed assessment in teacher evaluation. However, NWEA welcomes the opportunity to engage in or support serious investigations directed at assembling a strong validity argument.

**Requirement: Technical Characteristics 2.2**

2.2

Provide evidence of reliability, both for the total test and for any subtests for which scores are reported. Include estimates of error in measurement.

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

**Requirement: Technical Characteristics 2.2**

[Redacted content]

**Requirement: Technical Characteristics 2.3**

2.3  
Provide evidence that the assessment is appropriate for use with student subgroups, including English language learners and student with disabilities. Include documentation that the assessment does not exhibit bias toward any major subgroups (e.g., through an analysis of differential item functioning). In addition, provide a sensitivity review to demonstrate the assessment tasks and items are designed to be accessible and fair for all students.

The adaptive nature of MAP assessments make the tests appropriate for student subgroups with a wide range of skills and needs. When used consistently during instruction and assessment, NEWA approved accommodations pursuant to state and/or division policy are appropriate for an individual student.

**Use of MAP with Subgroups**

The adaptive nature of MAP assessments allows it to serve as an appropriate academic achievement measure for a broad range of students without modifications or accommodations. This includes many, but not all students with disabilities. The MAP system measures students by providing questions that match each student’s current level of challenge. This allows students with mild cognitive difficulties to be measured accurately by the tests. In addition, a wide variety of accommodations allow the tests to be given in a manner designed to preserve the accuracy of the measurement. A list of allowable accommodations is provided in Section 4.4.6.

Although students with disabilities such as specific learning disabled, mildly delayed, communication disorder, behavioral disorder, and mild motor impairment may take MAP assessments, collection of this level of detail about specific students is provided at the discretion of each school division. LEAs vary considerably in their willingness to provide such information routinely and in a standard format that would make it useful beyond the LEA providing it. For this reason, NWEA has not been in a position to initiate a comprehensive study of validity and reliability of MAP assessments for students with disabilities. Should the Department be interested in participating in such a study, NWEA would be pleased to work with division staff to identify methods and processes for collecting data to conduct an analysis on these special populations of

### **Requirement: Technical Characteristics 2.3**

students.

To establish validity and reliability evidence for the proposed MAP tests use for English Language Learners (ELL), several additional forms of data would be required. Since ELL students will vary in their exposure to and instruction in English, it would be important to verify differences in these areas. It is expected that this could be accomplished by accessing extant data sources such as student demographic and instructional records. To establish a reference level or progression of English language performance, scores from a single or defined set of common English language proficiency tests would be important. Ideally, the test(s) would have reliability estimates that meet conventionally accepted standards as well as validity evidence supporting their use in at least reading performance across difference ages and stages of development of English language proficiency. It would be desirable, though not absolutely essential, for the tests to be aligned to the WIDA English Language Proficiency Standard. The availability of a large representative sample of such data would allow: 1) validity evidence for the use of Virginia State-aligned MAP tests with ELL students, and 2) a determination to be made about the conditions and circumstances under which these MAP tests would not be valid.

#### **Differential Item Functioning Analysis**

The fundamental assumption underlying Item Response Theory is the probability of a correct response to a test item is a function of the item's difficulty and the person's ability. This function is expected to remain invariant to other personal characteristics unrelated to ability such as gender, ethnic group membership, family wealth, etc. Therefore, if two test takers with the same ability respond to the same test item, they are assumed to have an equal probability of answering the item correctly.

To view NWEA's DIF Analysis evidence that MAP does not exhibit bias towards major subgroups, refer to Appendix K.

NWEA takes steps to ensure tests are appropriate for students both prior to test construction and after test administration. Tests are comprised of items constructed in accordance with strict item writing guidelines that address both sensitivity and fairness. Sensitivity in this context means an awareness of the different things that can distract a student during assessment. Fairness in this context relates to giving each student equal opportunity to answer the item correctly based solely on their knowledge of the item content. Any sensitivity and fairness issues found in items are eliminated in revision prior to test construction.

Each item is evaluated against a set of criteria. An item is flagged if it:

- Requires prior knowledge other than the skill/concept being assessed
- Has cultural bias
- Has linguistic bias
- Has socio-economic bias
- Has religious bias
- Has geographic bias
- Has color-blind bias

**Requirement: Technical Characteristics 2.3**

- Has gender bias
- Inappropriately employs idiomatic English
- Offensively stereotypes a group of people
- Mentions body/weight issues
- Has inappropriate or sensitive topics (smoking, death, crime, violence, profanity, sex, etc.)
- Has other bias issues

**Bias and Sensitivity Review**

As part of the item development process, NWEA conducts a bias and sensitivity review. Sensitivity in this context means an awareness of the different aspects of an item or context that can distract a student during assessment. Fairness in this context relates to giving each student an equal opportunity to answer the item correctly based solely on his or her knowledge of the item content. A well-constructed item serves to activate and focus a student's thought process on the task presented in the item. A successful item is fair to all students. An item should NOT:

- Distract, upset, or confuse in any way.
- Require construct-irrelevant or specialized knowledge.
- Favor students from certain language communities.
- Favor students from certain cultural backgrounds.
- Favor students based on gender.
- Favor students based on social economic issues.
- Employ idiomatic or regional phrases and expressions.
- Stereotype certain groups of people or behaviors.
- Favor students from certain geographic regions.
- Favor students who have no visual impairments.

A hard and fast list of potentially distracting or upsetting material does not exist, but there are topics that are seldom appropriate for K-12 level assessments, such as sexuality, illegal substances, illegal activities, excessive violence, discriminatory descriptions, death, grieving, catastrophes, animal neglect or abuse, loss of family member, or weight and body issues. The best way to ensure that items are as fair and sensitive as possible is to review them specifically for sensitivity and fairness issues. Sensitivity and fairness reviews are incorporated into the Editorial review, Item Production review, and both sets of Content reviews.

**Requirement: Technical Characteristics 2.4**

2.4

Provide evidence that the assessment includes items of varying difficulty to ensure accurate measurement of student achievement across the ability continuum, including the tails of the score distribution.

**Requirement: Technical Characteristics 2.4**

MAP provides assessment items varying in difficulty with items appropriate for Grades 2-12 in mathematics, reading, and language usage (see Section 1.1.1). NWEA's psychometricians validate the item pools to ensure they are deep enough that a student taking a particular test will not see the same test item in a fourteen month time period, even if taking a test four times per year. MAP tests for each grade level are created from large item pools. For additional evidence of the varied difficulty of items that cover the expected ability distribution as well as sufficient item coverage, refer to Appendix G. In Appendix L, histograms of the mathematics, reading, and language usage item pools illustrate how items are distributed. Evidence of how these pools actually result in precise measurement across the ability continuum is presented in the SEM charts found in the last part of Appendix I.

**Section 3: Use of Assessment as a Measure of Growth**

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.1**

3.1

Provide evidence that the scores resulting from the assessment have been used as measures of growth by other local or state education agencies.

In August 2011, MAP for Mathematics, Reading, and Language Usage assessments were reviewed and approved by the New York State Education Department's list of approved growth assessments of the New York State Common Core Standards for use in teacher and principal evaluations. The list of approved assessments including MAP tests can be viewed on the following website: <http://usny.nysed.gov/rttt/teachers-leaders/assessments/home.html#assess>.

In April 2012, MAP for Mathematics, Reading, and Language Usage assessments were reviewed and approved by the Ohio Department of Education's list of approved assessments that can be used for a portion of teachers' and principals' evaluations. To be approved, NWEA submitted evidence that MAP assessments met the fundamental requirements for measuring student growth as identified by the Ohio Department of Education:

1. Be highly correlated with curricular objectives
2. Have enough "stretch" to measure the growth of both low-and high-achieving students
3. Meet appropriate standards of test reliability.

For additional evidence of use by local or state agencies, please see Section d. of the Proposal Narrative.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.2**

3.2

Describe the methodology used to measure growth. For example, does the assessment employ a vertical scale, use a computer-adaptive model to measure growth over time, or employ some other methodology. Does the methodology allow for the longitudinal measure of growth across academic years? What about the measurement of required growth on the proposed assessment to reach proficient on the statewide assessments (the Standards of Learning tests) in a specified amount of time? Include standard setting studies or other analyses conducted to establish measures of growth.

NWEA takes a very straightforward approach to estimating growth. Since the RIT scales are vertical (cross-grade) scales, growth is estimated directly based on difference (or change) between the status scores from two or more test events. These difference scores are qualified using the standard errors of the individual status scores involved in the comparison (specifically,  $\sqrt{SEM_{test1}^2 + SEM_{test2}^2}$ ). In order to take this approach, heavy reliance is placed on score precision and scale stability. Theoretically, the length of the interval between scores is only relevant when growth rate or comparison of growth (change) scores is of interest. Given this, growth can be computed and evaluated across two contiguous terms, across a school year, or longitudinally across multiple school years. To help inform the interpretation of observed growth, NWEA has developed a set of national norms to serve as points of reference. The latest (2011) norms continue a practice of publishing norms every three years, from 1996. The 2011 norms study is included on CD with this proposal.

The 2011 NWEA growth norms were developed using several methodological enhancements designed to improve their utility as a means of anticipating as well as evaluating student academic growth. These enhancements include:

- The use of empirical instructional weeks in estimating growth rates, which allowed the number of instructional weeks preceding the first test and varying test 1 – test 2 intervals to be used in normative growth projections
- Use of the first (pretest) score as a covariate in calculating normative growth projections
- Weighting scores by their standard errors in calculating normative growth projections
- Use of an empirically derived state-level School Challenge Index (based on NCES data) as a post-stratification weighting variable for both the growth and status norms.

As a result of these enhancements, student growth can be projected and evaluated by taking into account the student's initial RIT score, the time in the instructional year the first (e.g, fall) test is administered, the time interval between the two tests, and the standard error of the RIT score(s) involved. This provides a great deal of flexibility in how the RIT scale norms are used to help interpret growth as well as how growth standards could be approached and set. For example, a MAP scale alignment study to the Virginia State SOL tests was recently completed. Like other studies of this type, this study provides estimates of RIT scores associated with each cut score point at each grade level for mathematics and reading. Thus, even though a state does not use a cross-grade (vertical) scale, the alignment study will provide "bridges" from state test results to the (vertical) MAP scales. Using these bridges and the 2011 RIT scale growth norms as

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.2**

references, growth standards within and across grades could be meaningfully informed. Such standard setting could take advantage of the enhancements noted above to provide more precise estimates of growth for the purpose of projection or for evaluation. In addition, as a result of the 2011 RIT scale norms study, there has been renewed interest in refining the Hybrid Success Model (Kingsbury & McCall, 2005). The Hybrid Success Model provides an extension of shorter term growth estimates to those involving distal achievement targets (e.g., from spring of grade four to spring of grade seven). This model estimates the amount of growth required each year to meet the distal achievement target, given the student's current achievement level, the normative growth rate from that level, and the number of instructional weeks until the target will be assessed.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.3**

3.3

Describe the methodologies used to control item exposure so that the accuracy of students' scores is not impacted by multiple exposures to the same items.

Because MAP tests are adaptive, all test takers receive a customized form of the assessment each and every time they take a MAP assessment. NWEA controls the items each student sees by monitoring the specific items exposed to each student. Survey with Goals tests can be administered four times per year without the student seeing the same item until fourteen months have elapsed. In addition, the depth of the item pool makes it unlikely that students will encounter the same item again even after two years. Test algorithms assure items in the same content and difficulty category are chosen randomly, so that the students at the same ability level do not see the same items. NWEA continuously field tests new items each year to expand and maintain the integrity of the item banks. Should minor security breaches or items be exposed, NWEA will remove exposed items from the pool.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.4**

3.4

Describe the procedures used to validate the measures of growth.

The procedures to validate MAP as a growth measure have taken several forms. Formal investigations have been conducted into the construct validity and measurement invariance of MAP tests (Wang, McCall, Jiao & Harris, 2012<sup>1</sup>). In addition, the construct of longitudinal achievement using a vertical scale (RIT reading and mathematics scale, specifically) was validated using a latent growth modeling approach (Wang, Jiao & Zhang, 2012<sup>2</sup>). These studies provide strong support for the validity of the RIT scales as a platform to measure achievement growth. More traditional forms of validity such as concurrent validity have been difficult to initiate, primarily because educational entities commonly test students routinely only with a state sponsored accountability test. Such tests often assess at very gross time periods, making concurrent validity studies subject to key limitations.

Given that NWEA uses differences in (status) achievement estimates as estimates of achievement growth, it is imperative that status estimates (RIT scores) contain minimal measurement error.

SEM serves as a form of first level check on the confidence that can be placed in the score as a valid indicator of the student's ability. However, NWEA also monitors other indicators that contribute to making such judgments. Considerable attention is given to Individual Score Validity (ISV) by analyzing student effort demonstrated throughout the test event (Kingsbury & Wise, 2012<sup>3</sup>; Wise, 2012<sup>4</sup>; Wise, Kingsbury & Hauser, 2012<sup>5</sup>; Wise, Ma & Theaker, 2012<sup>6</sup>). Effort is operationalized using a set of heuristic flags triggered by combinations of rapid response times and incorrect answers in different portions of the test. Planning is currently underway for incorporating these ISV flags into production test scripts and reports. By helping to ensure test score validity, arguments for the validity of inferences that depend on growth are strengthened.

In addition to the measures NWEA takes to ensure score validity, schools, and school divisions are encouraged to undertake a number of procedures aimed at enhancing validity during test administration. The primary threats to validity that these procedures address include:

- Attempt to view the secured item pool prior to testing or to record items as testing occurs. The most common way this threat is realized is by teachers or administrators enrolling pseudo-students so that educators can take the test.
- Retesting students solely to improve a growth score. Since the primary purpose of interim testing is to provide information to help the teacher plan instruction, we believe very strongly that schools should not retest students with valid scores. Threats to validity emerge when retesting procedures are not consistently followed (for example if educators "cherry-pick" retesting to try and inflate their growth scores).
- Inappropriate help or accommodations during the test. Providing help to students that is not specifically authorized in an accommodation.
- Students who are not assigned to teachers within the accountability system.

We recommend that the following safeguards be in place in order to assure the integrity of the testing process, particularly when results are to be used for teacher evaluation. Most of these procedures are not unique to the MAP assessment; they are procedures we would recommend for administration of any test.

- Authorize only one person, preferably a non-teacher, in each school to enroll students (i.e., add student profile data) in the testing system. Most enrollment is done through the central office, but schools need the ability to enter new students into the system, sometimes at the last minute.
- Adopt a policy that prohibits enrolling non-students in the MAP system for the purpose of enabling teachers, parents, or others to take practice tests (thus exposing the item pool) and require that all students have a unique division- or state-issued ID. The division should periodically audit official enrollment records against the record of tested students to confirm that this policy is being followed.
- The MAP system provides data to schools about the validity of individual test events that can be used to ensure the measurement is accurate for each student. These data include test duration, percent correct (on adaptive tests, nearly all students should get approximately half the items right), and when required item-response data. We encourage schools to retest students when these data indicate that an administration is likely to have produced an invalid result. Adopt a policy to govern retesting of students in the MAP system, set specific empirical criteria for testing (for example, all students with test durations under fifteen minutes or percent correct under thirty-five percent should be retested), and require that

- policy be followed each term. Retesting requests should be approved by the principal, and the division should randomly audit records to assure that retesting policies are being followed.
- Require that accommodations be limited to students with Individual Education Plans and that accommodations allowed be limited to those specifically identified in the plan. Accommodations offered should be noted and approved by the principal for each student.
  - Adopt a policy that prohibits teachers from transcribing or saving items from the test and should communicate that policy to teachers.
  - MAP testing is normally conducted in a computer lab. We recommend that the test be proctored by both the teacher and a second person (normally the educator or assistant manning the lab). This discourages teachers offering inappropriate help to students. It also has several other benefits. First, the presence of the person responsible for the lab assures that someone in the room is available to troubleshoot problems if they occur. Second, it assures that the teacher has an opportunity to monitor testing conditions to ensure they are appropriate and that the test is properly administered.
  - If a computer lab is used for testing, we also encourage recording and preserving test activity on unobtrusive security cameras. This policy helps the division better assure test security, and also provides the testing procedures integrity in the event that a teacher is questioned on this.
  - The division should establish procedures to make sure assignments of students to teachers are accurately recorded in order to ensure the integrity of results. These procedures should require that all students are accounted for, and should require validation of teacher assignments by both the teacher and the school principal.

## Section 4: Test Administration Procedures

<b>Requirement: Section 4: Test Administration Procedures 4.1</b>
4.1 Describe the administration procedures necessary to produce growth scores. For example, is the assessment designed to be administered multiple times during the year or administered once in the fall and once in the spring?
MAP assessments are a measure of both status and growth. Beginning with the 2011 Norms Study, a student's achievement and growth can be referenced from an estimated growth trajectory the student shares with his/her academic peers (i.e., those students who performed equally well when tested during the same instructional week). Given the student's observed RIT score (and its accompanying SEM) from a test occurring at the specific instructional week, the student's normative performance on the test may be directly determined. The new approach may also be used to predict the student's performance if he/she had tested a week earlier, or several weeks later. As a result, performance may be evaluated at alternative and various points on the relevant instructional calendar and, in this sense, is less rigid than the more common test term window designations would suggest. The new procedure facilitates the determination of the Conditional Growth Index (CGI) which enables the user to directly determine the growth by the student's

**Requirement: Section 4: Test Administration Procedures 4.1**

second testing occasion relative to the student’s previous performance, and when compared with academic peers.

The depth of MAP item pools for Virginia are sufficient for four test administrations per year. NWEA provides four default testing window dates for the purpose of establishing data norms; however, Virginia LEAs will have the flexibility to select test term windows within the following default windows:

- Fall: August 15th through November 30th
- Winter: December 1st through February 28th
- Spring: March 1st through June 15th
- Summer: June 16th through August 14th

While flexible, it is important to establish beginning and ending dates for MAP testing for the purposes of data accuracy and division-level reports. The period of time defined by these dates is known as a test term window. The test term window is useful for setting an end date at which an LEA will declare the testing term complete and subsequently have the option to order division-level summary reports. NWEA reports will only display student results that occurred within a division's defined test term window.

NWEA offers Virginia divisions the following recommendations when selecting test window dates:

- Strive for a three week or less testing window to provide comparisons based on students that have received about the same amount of instruction.
- Keep test window dates consistent from one academic year to the next to allow for valid growth comparison from year to year.
- Establish test window dates using the following recommendations to utilize normative data populated into reports:

Time Category	NWEA Recommendation
Fall test window timing	Beginning of academic year (weeks 1-7)
Winter test window timing	Middle of academic year (weeks 15-21)
Spring test window timing	End of academic year (weeks 28-34)
Time between fall and spring testing	Roughly 32 weeks of instruction
Time between testing in consecutive fall or spring terms	Roughly 36 weeks of instruction

**Requirement: Section 4: Test Administration Procedures 4.2**

4.2

Describe any processes used for pre-identifying and/or registering students for testing. Include what data, including the State Testing Identifier, are collected for each student, how data are collected or transmitted, and how data are maintained and securely managed.

NWEA acknowledges the importance of rostering students prior to testing to ensure the right student receives the right assessment. As a result, NWEA offers comprehensive rostering processes and systems to address partner needs. To ease test administration, NWEA's rostering processes and systems are implemented to support finalizing registration in advance of test sessions. For those cases where a student must be registered immediately prior to testing, both MAP platforms provide a solution to roster individual students. For processing larger roster, or program files, NWEA proposes the following process and systems solution.

**Web-Based Platform**

Using the Administration and Reporting Center interface, division users will have control over the process of importing division-wide or school-wide data files to the web-based platform with student, teacher, class, school, and program information, including the State Testing Identifier. Division users assigned the Data Administrator role can upload complete or partial roster files and programs files to the system. Accommodation needs for individual students can be included in the roster files and will be tracked by the system.

The data files uploaded through the Administration and Reporting Center must adhere to the NWEA roster file and programs file templates; these templates define data fields, character limitations, and other restrictions based on business rules that enable the system to add and update existing information in the system with imported data. The roster files and program files templates are provided in Appendix M. Roster files and program files must be in the CSV (comma-separated values) file format to be imported into the web-based assessment system and can be accepted from any SIS system that allows for configurable data exports.

NWEA's web-based assessment system provides end users the flexibility to import all roster data at one time, or to submit partial data imports. The Assessment and Reporting Center interface accepts the following types of roster imports:

- **Instructor-Only Import** allows instructor access to tutorials, the student warm-up, user guides, and instructional resources, such as DesCartes: A Continuum of Learning, before classes begin. This is especially helpful for new instructor orientation. Divisions might also use this import type to update instructor names and e-mail addresses as needed.
- **Student-Only Import** is convenient for schools that need to prepare students for testing within a very short timeframe when instructor and class data is not yet available. This data can be augmented by updating individual student profiles or importing an updated roster file with more complete data. Note that students imported with minimum information required for testing (student name and grade) will not appear on reports until their profiles include all reporting attributes.
- **Complete Import** prepares student data so that students are ready for testing, organizes data for reporting, and provides instructor access to their students' reports.

Imported files undergo multiple data validation processes. If the system detects errors in the file

**Requirement: Section 4: Test Administration Procedures 4.2**

format (for example, the data file is not in CSV format) or inconsistencies in the data (for example, the same Student ID is assigned to two different students), the system generates an error report. The data administrator corrects the errors in the source system – typically the division’s Student Information System (SIS) – and submits an updated file. Updated roster files and programs files can be submitted via the Administration and Reporting Center interface as often as is needed.

**Client-Server Platform**

On the client-server platform, the student enrollment process is similar to that of the web-based platform described above, except:

- Instructor management is not supported; all user accounts are managed via NTE Admin
- The roster files, also in CSV file format, have slightly different data fields requirements
- The roster files are submitted via the NWEA Reports Site
- File validation is manually performed by the NWEA Technical Support team

**Data Transfer and Storage Security**

NWEA systems’ security architecture is based on the underlying security pillars of information security: Confidentiality, Integrity, and Availability. Confidentiality ensures that all NWEA system users can access only data that their permission level entitles them to access. Integrity ensures that unauthorized users are not allowed to tamper or alter any data they do not have access to. Availability ensures that NWEA systems and data are available to authorized users when they need them. The security architecture of NWEA systems is built to provide high security by employing security best principles.

NWEA systems transmit and receive data over the Internet using industry standard 256-bit secure sockets layer (ssl) using a 2048-bit certificate. This approach to secure data transfers is in compliance with the Family Education Rights and Privacy Act (“FERPA”) and Children’s Online Privacy Protection Act (“COPPA”).

**Requirement: Section 4: Test Administration Procedures 4.3**

4.3

Describe all materials needed for test administration and how school divisions will order and obtain sufficient quantities. Include details of test booklets and answer documents for paper/pencil testing (if applicable), test administration manuals, etc. If applicable, identify any test administration materials school divisions would be responsible for supplying locally (manipulatives, copies of test materials, etc).

MAP assessments are administered on a computer, typically in a school computer lab, therefore no printed test or administration materials are required. See Section 4.4.5 for the MAP system technical requirements. See Section 4.4.4 for a description of the resources NWEA provides MAP system users to help understand the administration of the assessments.

**Requirement: Section 4: Test Administration Procedures 4.4**

4.4

Provide examples of the test administration manuals to be used with the assessment(s).

NWEA provides multiple resources to help MAP users understand the general administration protocol that applies to all MAP assessments and test delivery platforms (client-server and web-based). These resources extend to general information about how MAP tests function and specific “how to” instructions for administering tests on each test delivery platform. Examples include:

- Guidelines for Teacher and Proctor Duties on Test Days, available on NWEA’s Partner Support website (<http://www.nwea.org/support/article/1042/teacher-and-proctor-duties-testing-days>) and included with this application as Appendix N.
- MAP Administration Characteristics of Tests, included in the NWEA Technical Manual and provided as Appendix O.
- MAP Proctor Handbook is available at <http://www.nwea.org/support/article/1150/proctor-handbook-map%C2%AE>.

Additional MAP administration documentation is available on the NWEA Partner Support website at <http://www.nwea.org/support/article-16>.

**Requirement: Section 4: Test Administration Procedures 4.5**

4.5

Describe all technology requirements related to school personnel managing the administration of tests and to students completing tests if assessments include technology-based delivery. Include the minimum and recommended hardware and software requirements and network requirements for test administration by school personnel and test delivery to students. Include how assessments are hosted (e.g., locally, vendor, 3<sup>rd</sup> party). Provide examples of user interfaces for test administration by school personnel and test delivery to students. Include descriptions or examples of test navigation and any test tools (e.g., calculator, ruler, highlighter) available to students for testing.

NWEA is one of the first providers of computer-based testing in the United States. Its long history in assessment technology puts it at the forefront of test design and test delivery systems. With this experience, NWEA can deliver to Virginia divisions a high-quality and reliable test with either a web-based or client-server platform. NWEA recognizes not all of Virginia’s schools are equipped with the Internet access required to use a web-based product, so NWEA also offers our client-server platform. These platforms have successfully administered over 262,800 tests in the Commonwealth of Virginia during the 2011-2012 school year.

**NWEA’s Web-Based Platform**

The web-based platform is an end-to-end online testing solution with components that comply with the highest standards for data security, transfer speed, and disaster recovery, while providing outstanding up-time during scheduled system availability. The web-based platform is accessible to any computer inside or outside of the division that meets the hardware and software requirements provided in Appendix P. Key components of NWEA’s web-based system are:

#### Requirement: Section 4: Test Administration Procedures 4.5

- **Online Administration and Reporting Center**, providing authorized users role-based access to administration functionality such as: user, student, organization, program, and test data management. Users will also access the Administration and Reporting Center to create and manage student test sessions, access reports, and instructional resources.
- **The Student Testing Center**, a web-based test taking environment where students log in using a standard web browser to take tests that have been scheduled by proctors.
- **Computer-adaptive tests** designed to measure achievement of students as they grow over time. The tests commonly consist of more than forty questions so the system can provide sub-scores that represent the content goals and an overall score for each student.

NWEA plans to add a new capability to our web-based platform to allow for lower grade Internet connections, expected to be available by July 2014. Until the new capability is available, NWEA recommends divisions and schools with lower-grade Internet connections use the client-server platform as an interim solution. NWEA will carefully monitor the progress of high bandwidth internet service as this capacity grows statewide. We will work with the Virginia DOE and the divisions to determine the optimal process to move divisions to our web-based solution.

#### NWEA's Client-Server Platform

The client-server platform is an end-to-end testing solution for administering tests on the TestTaker Client workstation without Internet connection. Four main components comprise the client-server assessment system:

- **Network Test Environment (NTE) folder(s)**. The MAP Assessments are designed to be administered in a local network environment. The NTE is a folder located on the local network that contains agency data, student data, and test package data.
- **NTE Administration Tool (NTE Admin)**. NTE Admin is an installed application used for managing a division's NTE. This program can be used to perform tasks such as adding students to an NTE on a student-by-student basis, moving students from one NTE to another or uploading test results. To use the NTE Admin, you will need an Internet connection and Microsoft .NET Framework 4.0. Since NTE Admin is only used for administrative functions and not for the actual test administration, it is not necessary to have NTE Admin installed on every computer used for test administration. NWEA recommends at least one computer in each school have NTE Admin installed, and that these computers have a high-speed Internet connection.
- **TestTaker Client**. TestTaker is the software application that students use to take NWEA assessments. Proctors will log in to TestTaker to select the test to be administered and the student to whom the test will be administered. Proctors also have access to the Proctor Administration Menu, where they can pause or terminate a test if needed. Running TestTaker requires network access to the NTE server but does not require an Internet connection.
- **NWEA Reports Site**. The NWEA Reports Site is a secure website where division assessment coordinators, school administrators, and teachers can log in to see their students' test results. The Reports Site also includes administrative tools to help the division assessment coordinator manage the testing season, including a secure upload page for Class Roster Files, a page for declaring division testing complete, and a report order page for school and division

**Requirement: Section 4: Test Administration Procedures 4.5**

summary reports.

The client-server user interface for test administration and delivery by school personnel is nearly parallel to that of the web-based platform.

Both the web-based and client-server assessment systems are deployed and operated independently for each Virginia division, requiring no state-level mediation. Technical requirements for both MAP platforms are included as Appendix P.

**Navigation Tools**

MAP for Mathematics assessments include a four-function calculator.

**Requirement: Section 4: Test Administration Procedures 4.6**

4.6

Describe accommodations available to students with disabilities and limited English proficient students. Include procedures related to the provision of accommodations to eligible students.

The adaptive nature of MAP assessments make them appropriate for many students in particular subgroups (i.e., students with disabilities, limited English proficient students, etc.). We recognize that student needs and skills within these groups are not homogeneous. In view of this, we focus accommodations on skills and practices used consistently during instruction and assessment. NWEA-approved accommodations pursuant to state and/or division policy are appropriate for an individual student. The following guidance is available when administering MAP to subgroups requiring accommodations:

**General Guidelines for Accommodations**

- Responsibilities: Scribes, page-turners, educational assistants, and other people supporting a student's test must be neutral in responding to the student during test administration. Assistance in test administration must not lead a student to the correct answer. The student's response must accurately represent the student's own choice.
- Reading Tests: Do not read aloud any portion of the reading test to any student.
- Symbols: Do not pronounce or explain math or science symbols.
- Definitions: Do not define any words within test questions.
- Interpreting Results: While many accommodations are recognized, NWEA has not performed studies to confirm how their use affects assessment scores. By definition, accommodations should allow more precise and more valid estimates of true scores. That is, accommodations are intended to "level the playing field" for students with known disabilities. They should not advantage any student relative to any other student at the same true ability level. However, for certain presentation accommodations noted below (2 and 3), we cannot state with a high level of certainty how validity of test scores is affected due to the potential change in the construct being measured. This qualification should be considered when interpreting test scores or using them to make important educational decisions.

**Requirement: Section 4: Test Administration Procedures 4.6**

**Presentation Accommodations**

1. Simplify or clarify directions
2. Native language translation (oral or signing) of test directions  
NOTE: See “Interpreting Results” under the general guidelines above.
3. Native language translation (oral or signing) of test questions (not answer options) for math, science, or language usage tests  
NOTE: See “Interpreting Results” under the general guidelines above.
4. Read or reread aloud the test directions
5. Read or reread aloud the test questions (not answer options) for math, science, or language usage tests only  
NOTE: See “Symbols” under the general guidelines above.
6. Use visual magnification devices or software (for example, MAP is compatible with ZoomText or MAGic)
7. Use auditory amplification devices, noise buffers, or software
8. Use masks to block portion of screen; for example, the student may use a sticky note, index card, or a blank sheet of paper to move down the screen as he or she is reading

**Response Accommodations**

1. Assign scribe to record responses
2. Dictate responses to a scribe
3. Point to responses for a scribe
4. Respond in native language

**Setting Accommodations**

1. Test an individual student in a separate setting
2. Test a small group of students in a separate setting; for example, in a Title I room or counselor’s office
3. Minimize distractions; for example, use a study carrel

**Timing/Schedule Accommodations**

1. Administer test over multiple sessions in a day
2. Administer test over a number of days
3. Administer test at a particular time of day
4. Allow flexible schedule
5. Extend time allotted by proctor or test administrator (All MAP assessments are designed to be untimed.)
6. Allow stop-the-clock breaks
7. Administer at time of day most beneficial to student

**Requirement: Section 4: Test Administration Procedures 4.6**

8. Offer breaks

**Materials or Devices Accommodations**

1. Provide scratch paper
2. Availability of a calculator (when appropriate, it is provided on the screen)

**Miscellaneous Accommodations**

1. Provide drink during testing
2. Provide snack during testing

In addition to these accommodations, MAP is compatible with third party software to provide capabilities to address the needs of visually impaired students including magnification, highlighting, and color contrast. NWEA will explore MAP compatibility with third-party software to accommodate the needs of blind students by evaluating third party software compatibility of text to speech and electronic Braille.

The following table identifies which English Language Learner (ELL) accommodations to the MAP assessment are available within the assessment:

- Read directions orally in English
- Allow flexible schedule – test in sessions, over several days
- Extended time
- Provide scratch paper/allow to mark test booklet
- Administer in a small group
- Administer individually

**Requirement: Section 4: Test Administration Procedures 4.7**

4.7

Describe procedures for completed student tests to be submitted for scoring and reporting purposes.

Since 2000, all computer-administered NWEA assessments have been computer-scored by the MAP assessment system as tests are administered. Computer scoring allows NWEA to provide preliminary testing results for the student and teacher at the end of each test to identify students needing intervention in a timely manner and to accurately point to instructional learning objects. Scoring of MAP assessments is based on Item Response Theory (IRT). Student scores are calculated, not only by employing an algorithm that takes into account the students' performance on the test items, but also the difficulty of the items each student is administered. NWEA's system has been rigorously tested to ensure scores rendered from the system are calculated and displayed correctly.

## Section 5 : Scoring and Reporting

### Requirement: Section 5: Scoring and Reporting 5.1

#### 5.1

Describe scoring procedures for all item types and test forms administered, including implemented quality control measures.

MAP assessments are adaptive and employ a common scoring algorithm. During the assessment, a Bayesian scoring algorithm is used to inform item selection. Bayesian scoring for item selection prevents the artificially dramatic fluctuations in student achievement at the beginning of the test that can occur with other scoring algorithms.

Although the Bayesian scoring works well as a procedure for selecting items during test administration, Bayesian scores are not appropriate for the calculation of final student achievement scores. This is because Bayesian scoring uses information other than the student's responses to questions, such as past performance, to calculate the achievement estimate. Since only the student's performance on the day he/she takes a test should be used to give the student's current score, a maximum-likelihood algorithm is used to calculate a student's actual score at the completion of the test.

NWEA uses the Rasch Item Response Theory model to create its vertical scales, called RIT scales. MAP results, reported as RIT scores, relate directly to the curriculum scale in each subject area.

At inception, NWEA adopted Item Response Theory in general and the Rasch model in particular to create its highly stable vertical RIT scales. There is one RIT scale each for mathematics, reading, and language usage. Using the RIT scale to report test results makes it possible to follow a student's educational growth from testing season to testing season and year to year. Please note that each subject area has a unique alignment to the RIT scale; as a result, scores between subjects are not equivalent.

By using Item Response Theory to create the scales and anchoring item difficulty estimates to them, the RIT scales are comparable from one set of items to another, and from one set of examinees to another. This enables comparisons of the scores from different students, or from the same student at different times, even though different sets of test items are administered. This also allows longitudinal comparison of student performance to be made.

#### **QA Process**

NWEA's dedicated QA staff work together with NWEA's research team to perform a comprehensive review of all reporting data for accuracy and quality. This includes the validation of all business rules and formulas applied when generating results reports for both standard reports provided via the assessment platform and all custom reports or data extracts provided.

**Requirement: Section 5: Scoring and Reporting 5.2**

5.2

Describe the type of reporting provided (e.g, static and/or dynamic, electronic and/or paper-based, item-level, strand-level, and/or test-level scoring). Include approximate timelines for score reports to be available to divisions, how score reports will be accessed and/or obtained, and samples of student, class, school, and division score reports and sample record layouts for electronic data files.

NWEA assessments provide a comprehensive set of static electronic reports, instructional resources, data tools and comparative information to gain insight into student growth and achievement at the individual, classroom, school, grade and division level. System reports and instructional resources are student-centric, research-based, and data-driven. They provide data to inform instruction, evaluate programs, justify budget initiatives, and make key decisions in the classroom, school, and division levels.

These analytical tools include the following reports:

- Student-level reports including the Student Progress Report and the Student Goal-Setting Worksheet provide current and historical student performance. This includes growth and growth projections as well as normative data and Lexile<sup>®</sup> scores to assist students, teachers, and parents select suitable reading texts. Student Progress Reports can be printed by the educator and shared with parents to create a parent/teacher dialog around their student's learning, performance, and growth. To help guide parents through the various components of this report, a special one page annotated version of the report is provided as an accompaniment. The Student Goal Setting Worksheet is especially helpful in allowing parents to see the growth potential of their child.
- Class and teacher reports provide an overview of class performance and detailed information about each student, including: assessment and goal area performance, norm percentile rank, growth, and predictive performance on the summative assessments once a scale alignment study is completed. Teachers can use these reports to differentiate instruction and identify areas of strength and weakness at a group level.
- School-level reports provide information to help administrators evaluate programs and foster collaboration.
- Grade-level reports provide summary information to help school administrators assess trends, identify areas of strength and weakness, and percentage of students meeting their growth targets
- Division-level reports which help partners assess performance trends by grade and school. These reports present longitudinal data for a particular school and are valuable resources in planning and monitoring school improvement plans.

NWEA uses a user-centric approach to report design. Throughout the product lifecycle, usability testing is incorporated to understand how a teacher or administrator intends to use the reports, and incorporates this mindset in the development of report designs that are further enforced by the experience of NWEA's research teams and education experts. This user-centric approach provides greater benefit to partners through limiting the amount of time required by the user to learn the content of the reports. The information contained within our reports is intended to aid in

**Requirement: Section 5: Scoring and Reporting 5.2**

swift and decisive action plans.

Longitudinal Results: NWEA collects and stores student MAP assessment results across time, providing educators with longitudinal data to compare student growth over time. Because MAP results are reported on vertical scales called the RIT scales, it is possible to compare individual students' results over time and compare sets of results. Furthermore, educators can access reports for the current academic year and the previous academic year from NWEA's Reports Site (see page 5 of Appendix Q for a sample ).

DesCartes: A Continuum of Learning: For students in grades 2-12, NWEA offers DesCartes: A Continuum of Learning. DesCartes translates student goal scores into relevant skills and concepts by achievement level and are aligned to Virginia's Goal Structures created from the Virginia SOL. DesCartes can be used to form and implement flexible grouping strategies, and identify and monitor student needs according to student progress.

**Availability and Format of Reports**

The MAP assessment system scores each test as it is administered and provides preliminary results for the student and teacher at the end of each test. Reports showing test results are available to teachers and administrations via login at the web-based reporting interface. For the Web-based MAP system, reports are accessed via the MAP Administration and Reporting Center. Please refer to the table below for complete information regarding reports availability. Reports can either be viewed online and/or using a PDF viewer, such as Adobe Reader®. From the PDF viewer reports can be viewed, saved, or printed.

Available Immediately	Available within 24 hours*
<ul style="list-style-type: none"> <li>▪ Class Breakdown by RIT Report</li> <li>▪ Class Breakdown by Goal Report</li> <li>▪ Class by Projected Proficiency</li> <li>▪ Descartes: A Continuum of Learning</li> <li>▪ Primary Grades Instructional Data</li> <li>▪ Student Goal Setting Worksheet</li> </ul>	<ul style="list-style-type: none"> <li>▪ District (Division) Summary Report</li> <li>▪ Student Progress Report</li> <li>▪ Class Report</li> <li>▪ Grade Report</li> <li>▪ Achievement Status and Growth Reports</li> <li>▪ Student Growth Summary Report</li> <li>▪ Projected Proficiency Summary Report</li> </ul>

\*These reports may be available immediately depending on the amount of data.

In the client-server MAP system, reports are accessed via login to NWEA's web-based Reports Site. Please refer to the table below for information about reports availability timeframe and format.

**Requirement: Section 5: Scoring and Reporting 5.2**

Available Reports	Timeframe	Format
Teacher- and class-level reports	The day after test results are uploaded to NWEA	PDF*
Grade-, school-, and district-level (division-level) reports	Within 24-72 hours after requesting reports	Crystal Reports*, tab-delimited (Data Text File), and CSV (Comprehensive Data Files)
Dynamic Reports showing class-, school-, grade-, and district-level (division-level) growth and proficiency data	One week after test results are uploaded to NWEA	Online display; can be exported to PDF

**Report Samples**

Samples of student, class, school, and division reports are included in Appendix Q, Annotated Reports. In addition to sample reports, Appendix Q provides guidance on how to interpret these reports. The annotated reports display sample versions along with helpful explanations of what each of the different report components mean. NWEA partners use the Annotated Reports to gain better understanding for how to interpret assessment scores. Additionally, NWEA offers onsite workshops and online training specifically designed to help educators interpret results and adjust instructional practices to meet their students’ academic needs. See Proposal Narrative, Section a. for information about NWEA’s online support and professional development workshops.

**Requirement: Section 5: Scoring and Reporting 5.3**

5.3  
 Describe all data tools available to school division staff for the analysis of data and the creation of customized reports.

NWEA is committed to providing accurate and timely results to Virginia educators. In addition to offering the aforementioned menu of reports, NWEA also makes available two additional options for obtaining data: raw data files and custom reporting.

**Raw Data Files**

For divisions, the MAP system generates a Comprehensive Data File that contains assessment results, student data, class assignments, and special program (sub-group) designations in a flat text format (common-separated values (CSV)) for further analysis and use in external data or reporting systems, i.e. data warehouses. See Appendix M for file format sample.

**Optional Division Custom Reporting**

If Virginia divisions require custom reporting capabilities, NWEA is able to deliver sophisticated custom reports designed alongside our partners at a negotiated rate.

- 
- <sup>1</sup> Wang, S., McCall, M., Jiao, H., & Harris, G. (April, 2012). Construct Validity and Measurement Invariance of Computerized Adaptive Testing: Application to Measures of Academic Progress (MAP) Using Confirmatory Factor Analysis. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.
- <sup>2</sup> Wang, S., Jiao, H., & Zhang, L. (2012, April). Validation of Longitudinal Achievement Constructs of Vertically Scaled Computerized Adaptive Tests: A Multiple Indicator Latent Growth Modeling Approach. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.
- <sup>3</sup> Kingsbury, G. G., & Wise, S. L. (in press). Turning the page: How smarter testing, vertical scales, and understanding of student engagement may improve our tests. In R. W. Lissitz and H. Jiao (Eds.) *Computers and their impact on state assessment: Recent history and predictions for the future*. Charlotte, NC: Information Age Publishing, Inc.
- <sup>4</sup> Wise, S. L. (2012). *The utility of adaptive testing in addressing the problem of unmotivated examinees*. Manuscript submitted for publication.
- <sup>5</sup> Wise, S. L., Kingsbury, G. G., & Hauser, C. (2012). *How do I know that this score is valid? The case for assessing individual score validity*. Manuscript submitted for publication.
- <sup>6</sup> Wise, S. L., Ma, L., & Theaker, R. A. (2012, May). *Identifying non-effortful student behavior on adaptive tests: Implications for test fraud detection*. Paper presented at the Conference on the Statistical Detection of Potential Test Fraud, Lawrence, Kansas.

**Offeror Name:** Northwest Evaluation Association (NWEA)  
**Proposed Assessment Name:** MAP for Science Assessments  
**Content Area(s) and Grade Level(s) Assessed:** Science for Grades 3-10

## Section 1: Overview of Tests

### Requirement: Section 1: Overview of Tests 1.1

#### 1.1

Describe the specific grade(s) and subject area(s) covered by each assessment and provide an overview of the content and skills measured. Include the types of test items used, the mode(s) of delivery, the availability of equivalent forms, including short forms or screeners (if available) and a test blueprint for each test being proposed.

MAP for Science assessments are computer-adaptive tests containing selected-response items. It is appropriate for students in grades 3-10, and is aligned to the Virginia Standards of Learning (Virginia SOL). MAP for Science Survey with Goals tests can be administered up to three times per year. Survey with Goals tests typically consist of more than forty questions and provide an overall RIT score for the subject as well as sub-scores in each of the goal performance areas. The accuracy of these tests allows the results to be used to measure student growth as well as current achievement level.

One characteristic of NWEA's MAP for Science assessment is that the test adapts item by item to give each student a unique testing experience. Science assessments:

- Provide an overall RIT (see Section 2.2.2 for a complete description of NWEA's RIT scales) score along with sub-scores in each of the goal performance areas.
- Are adaptive tests designed to measure achievement of students as they grow through the grades. The accuracy of these tests allows the results to be used to measure student growth as well as current achievement level status.

MAP for Science assesses a student's understanding of specific science concepts within the three major domains of science: life sciences, earth and space sciences, and physical sciences. The content and skills covered by NWEA's Virginia State-aligned MAP for Science assessments are demonstrated in Appendix R, which contains NWEA's Goal Structure Charts. Appendix R also includes an explanation of how the structure of the Virginia SOL was addressed in the development of the Virginia State-aligned MAP for Science assessments. See Appendix S for the MAP for Science Test Blueprint.

### Requirement: Section 1: Overview of Tests 1.2

#### 1.2

Provide evidence of alignment of test items to the Virginia Standards of Learning (SOL) for existing assessments. For assessments developed in response to the RFP, provide a plan for assuring the alignment of test items to the SOL.

**Requirement: Section 1: Overview of Tests 1.2**

The MAP assessments and associated reports are based upon test-specific goal structures created by NWEA Content Specialists. These Content Specialists are content area experts with five or more years of classroom teaching experience. They have an expert level of knowledge of writing and reviewing items for classroom, district, interim, and state assessments; have an expert knowledge of national trends and initiatives in assessment. Additionally, they have experience creating test structures aligned to state standards and selecting item pools for those assessments. In the alignment process, NWEA Content Specialists reviewed the Virginia SOL standards carefully and organized them into reporting frameworks (Goal Structures) for MAP tests. A goal structure was created by combining Virginia's content standards and grade-level expectations across grade levels to create a two-tier framework consisting of goals and sub-goals. Once the goal structures were finalized, NWEA Content Specialists aligned items from NWEA's item bank to each subgoal. This ensured the item pools for the Virginia assessments contain only topics within the Virginia SOL standards.



**Section 2: Technical Characteristics**

**Requirement: Section 2: Technical Characteristics 2.1**

2.1

Provide evidence of content, construct, concurrent, and predictive validity as appropriate. Include validity evidence that supports the use of scores from the proposed assessment in teacher evaluation, addressing specifically the validity of using assessment results to support inferences about effectiveness of teacher in producing growth in student performance (if available).

**Requirement: Section 2: Technical Characteristics 2.1**

Many types of information can be used as validity evidence. This information ranges, for example, from the adequacy and coverage of a test’s content, to its ability to yield scores predictive of a status in some area, to its ability to draw accurate inferences about a test taker’s status with respect to a construct, to its ability to allow generalizations from test performance within a domain to like performance in the same domain. Some common forms of validity evidence can be offered to support broad validity arguments. Appendix U includes information about content, concurrent, predictive, and criterion-related validity. The vast preponderance of this evidence comes from the relationships of MAP test scores to state content-aligned accountability test scores. These come in several forms including:

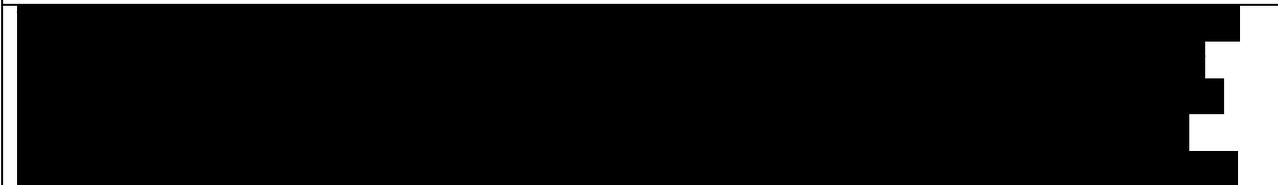
- the test content;
- the concurrent performance of students on MAP tests with their performance on state tests given for accountability purposes;
- the predictive relationship between students’ performance on MAP tests with their performance, two testing terms later, on state accountability tests; and
- the relationship between students’ performance on MAP tests and their nominal status relative to criteria defined by their state’s achievement standards.

MAP assessments are first and foremost, aimed at providing estimates of individual students’ academic achievement that minimize measurement error, whether that error is directional in some form of bias or the result of unstable scales. NWEA warrants MAP tests to be valid for uses related to these characteristics. Such uses may entail, for example, making decisions about a student’s achievement status, or about a student’s change(s) in achievement status (growth), or understanding how a student’s achievement status and growth compare to carefully constructed national norms. We have no conclusive or even systematically collected evidence about what antecedents contribute to or detract from a student’s achievement estimates. We suspect that such relationships will be shown to be multivariate and highly multidimensional. Whatever the structure of these relationships, strong and convincing evidence about them will be fundamental to any argument that attributes a particular outcome to a particular antecedent or set of antecedents. Due to the absence of such basic evidence, we have not constructed validity evidence that supports the use of scores from the proposed assessment in teacher evaluation. However, NWEA welcomes the opportunity to engage in or support serious investigations directed at assembling a strong validity argument.

**Requirement: Technical Characteristics 2.2**

2.2

Provide evidence of reliability, both for the total test and for any subtests for which scores are reported. Include estimates of error in measurement.



**Requirement: Technical Characteristics 2.2**

[Redacted text block]

- [Redacted text block]

**Requirement: Technical Characteristics 2.3**

2.3

Provide evidence that the assessment is appropriate for use with student subgroups, including English language learners and student with disabilities. Include documentation that the assessment does not exhibit bias toward any major subgroups (e.g., through an analysis of differential item functioning). In addition, provide a sensitivity review to demonstrate the assessment tasks and items are designed to be accessible and fair for all students.

The adaptive nature of MAP assessments make the tests appropriate for student subgroups with a wide range of skills and needs. When used consistently during instruction and assessment, NEWA approved accommodations pursuant to state and/or division policy are appropriate for an individual student.

**Use of MAP with Subgroups**

The adaptive nature of MAP assessments allows it to serve as an appropriate academic achievement measure for a broad range of students without modifications or accommodations. This includes many, but not all students with disabilities. The MAP system measures students by providing questions that match each student's current level of challenge. This allows students with mild cognitive difficulties to be measured accurately by the tests.

In addition, a wide variety of accommodations allow the tests to be given in a manner designed to preserve the accuracy of the measurement. A list of allowable accommodations is provided in Section 4.4.6.

Although students with disabilities such as specific learning disabled, mildly delayed, communication disorder, behavioral disorder, and mild motor impairment may take MAP assessments, collection of this level of detail about specific students is provided at the discretion of each school division. LEAs vary considerably in their willingness to provide such information routinely and in a standard format that would make it useful beyond the LEA providing it. For this reason, NWEA has not been in a position to initiate a comprehensive study of validity and reliability of MAP assessments for students with disabilities. Should the Department be interested in participating in such a study, NWEA would be pleased to work with division staff to identify methods and processes for collecting data to conduct an analysis on these special populations of students.

To establish validity and reliability evidence for the proposed MAP tests use for English Language Learners (ELL), several additional forms of data would be required. Since ELL students will vary in their exposure to and instruction in English, it would be important to verify differences in these areas. It is expected that this could be accomplished by accessing extant data sources such as student demographic and instructional records. To establish a reference level or progression of English language performance, scores from a single or defined set of common English language proficiency tests would be important. Ideally, the test(s) would have reliability estimates that meet conventionally accepted standards as well as validity evidence supporting their use in at least reading performance across difference ages and stages of development of English language proficiency. It would be desirable, though not absolutely essential, for the tests to be aligned to the WIDA English Language Proficiency Standard. The availability of a large

### **Requirement: Technical Characteristics 2.3**

representative sample of such data would allow: 1) validity evidence for the use of Virginia State-aligned MAP tests with ELL students, and 2) a determination to be made about the conditions and circumstances under which these MAP tests would not be valid.

#### **Differential Item Functioning Analysis**

To date, NWEA has not completed a DIF Analysis of MAP for Science tests. However, if sufficient data are available, NWEA will include a DIF Analysis for science in its next release of its Technical Manual. NWEA does not expect the proportions of items identified as having DIF to exceed the proportions observed for mathematics, reading, and language usage.

Each item is evaluated against a set of criteria. An item is flagged if it:

- Requires prior knowledge other than the skill/concept being assessed
- Has cultural bias
- Has linguistic bias
- Has socio-economic bias
- Has religious bias
- Has geographic bias
- Has color-blind bias
- Has gender bias
- Inappropriately employs idiomatic English
- Offensively stereotypes a group of people
- Mentions body/weight issues
- Has inappropriate or sensitive topics (smoking, death, crime, violence, profanity, sex, etc.)
  - Has other bias issues

Refer to Appendix K for an overview of NWEA's DIF analysis findings for MAP (mathematics, reading, and language usage).

#### **Bias and Sensitivity Review**

As part of the item development process, NWEA conducts a bias and sensitivity review. Sensitivity in this context means an awareness of the different aspects of an item or context that can distract a student during assessment. Fairness in this context relates to giving each student an equal opportunity to answer the item correctly based solely on his or her knowledge of the item content. A well-constructed item serves to activate and focus a student's thought process on the task presented in the item. A successful item is fair to all students. An item should NOT:

- Distract, upset, or confuse in any way.
- Require construct-irrelevant or specialized knowledge.
- Favor students from certain language communities.
- Favor students from certain cultural backgrounds.

**Requirement: Technical Characteristics 2.3**

- Favor students based on gender.
- Favor students based on social economic issues.
- Employ idiomatic or regional phrases and expressions.
- Stereotype certain groups of people or behaviors.
- Favor students from certain geographic regions.
- Favor students who have no visual impairments.

A hard and fast list of potentially distracting or upsetting material does not exist, but there are topics that are seldom appropriate for K-12 level assessments, such as sexuality, illegal substances, illegal activities, excessive violence, discriminatory descriptions, death, grieving, catastrophes, animal neglect or abuse, loss of family member, or weight and body issues. The best way to ensure that items are as fair and sensitive as possible is to review them specifically for sensitivity and fairness issues. Sensitivity and fairness reviews are incorporated into the Editorial review, Item Production review, and both sets of Content reviews.

**Requirement: Technical Characteristics 2.4**

2.4

Provide evidence that the assessment includes items of varying difficulty to ensure accurate measurement of student achievement across the ability continuum, including the tails of the score distribution.

MAP provides assessment items varying in difficulty with items appropriate for grades 3-10 in science (see Section 1.1.1). NWEA's psychometricians validate the item pools to ensure they are deep enough that a student taking a particular test will not see the same test item in a fourteen month time period, even if taking a test three times per year. The MAP for Science assessments are created from large item pools.

For additional evidence of the varied difficulty of items that cover the expected ability distribution as well as sufficient item coverage, refer to Appendix T. In Appendix W, histograms of science item pools illustrate how items are distributed.

**Section 3: Use of Assessment as a Measure of Growth**

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.1**

3.1

Provide evidence that the scores resulting from the assessment have been used as measures of growth by other local or state education agencies.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.1**

In August 2011, MAP for Science assessments were reviewed and approved by the New York State Education Department's list of approved growth assessments of the New York State Common Core Standards for use in teacher and principal evaluations. The list of approved assessments including MAP tests can be viewed on the following website: <http://usny.nysed.gov/rttt/teachers-leaders/assessments/home.html#assess>.

In April 2012, MAP for Science assessments were reviewed and approved by the Ohio Department of Education's list of approved assessments that can be used for a portion of teachers' and principals' evaluations. To be approved, NWEA submitted evidence that MAP assessments met the fundamental requirements for measuring student growth as identified by the Ohio Department of Education:

1. Be highly correlated with curricular objectives
2. Have enough "stretch" to measure the growth of both low-and high-achieving students
3. Meet appropriate standards of test reliability.

For additional evidence of use by local or state agencies, please see Section d. of the Proposal Narrative.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.2**

3.2

Describe the methodology used to measure growth. For example, does the assessment employ a vertical scale, use a computer-adaptive model to measure growth over time, or employ some other methodology. Does the methodology allow for the longitudinal measure of growth across academic years? What about the measurement of required growth on the proposed assessment to reach proficient on the statewide assessments (the Standards of Learning tests) in a specified amount of time? Include standard setting studies or other analyses conducted to establish measures of growth.

NWEA takes a very straightforward approach to estimating growth. Since the RIT scales are vertical (cross-grade) scales, growth is estimated directly based on difference (or change) between the status scores from two or more test events. These difference scores are qualified using the standard errors of the individual status scores involved in the comparison (specifically,  $\sqrt{SEM_{test1}^2 + SEM_{test2}^2}$ ). In order to take this approach, heavy reliance is placed on score precision and scale stability. Theoretically, the length of the interval between scores is only relevant when growth rate or comparison of growth (change) scores is of interest. Given this, growth can be computed and evaluated across two contiguous terms, across a school year, or longitudinally across multiple school years. To help inform the interpretation of observed growth, NWEA has developed a set of national norms to serve as points of reference. The latest (2011) norms continue a practice of publishing norms every three years, from 1996. The 2011 norms study is included on CD with this proposal.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.2**

The 2011 NWEA growth norms were developed using several methodological enhancements designed to improve their utility as a means of anticipating as well as evaluating student academic growth. These enhancements include:

- The use of empirical instructional weeks in estimating growth rates, which allowed the number of instructional weeks preceding the first test and varying test 1 – test 2 intervals to be used in normative growth projections
- Use of the first (pretest) score as a covariate in calculating normative growth projections
- Weighting scores by their standard errors in calculating normative growth projections
- Use of an empirically derived state-level School Challenge Index (based on NCES data) as a post-stratification weighting variable for both the growth and status norms.

As a result of these enhancements, student growth can be projected and evaluated by taking into account the student’s initial RIT score, the time in the instructional year the first (e.g, fall) test is administered, the time interval between the two tests, and the standard error of the RIT score(s) involved. This provides a great deal of flexibility in how the RIT scale norms are used to help interpret growth as well as how growth standards could be approached and set. For example, a MAP scale alignment study to the Virginia State SOL tests was recently completed. Like other studies of this type, this study provides estimates of RIT scores associated with each cut score point at each grade level for mathematics and reading. Thus, even though a state does not use a cross-grade (vertical) scale, the alignment study will provide “bridges” from state test results to the (vertical) MAP scales. Using these bridges and the 2011 RIT scale growth norms as references, growth standards within and across grades could be meaningfully informed. Such standard setting could take advantage of the enhancements noted above to provide more precise estimates of growth for the purpose of projection or for evaluation. In addition, as a result of the 2011 RIT scale norms study, there has been renewed interest in refining the Hybrid Success Model (Kingsbury & McCall, 2005). The Hybrid Success Model provides an extension of shorter term growth estimates to those involving distal achievement targets (e.g., from spring of grade four to spring of grade seven). This model estimates the amount of growth required each year to meet the distal achievement target, given the student’s current achievement level, the normative growth rate from that level, and the number of instructional weeks until the target will be assessed.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.3**

3.3

Describe the methodologies used to control item exposure so that the accuracy of students’ scores is not impacted by multiple exposures to the same items.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.3**

Because MAP tests are adaptive, all test takers receive a customized form of the assessment each and every time they take a MAP assessment. NWEA controls the items each student sees by monitoring the specific items exposed to each student. Survey with Goals tests can be administered three times per year without the student seeing the same item until fourteen months have elapsed. In addition, the depth of the item pool makes it unlikely that students will encounter the same item again even after two years. Test algorithms assure items in the same content and difficulty category are chosen randomly, so that the students at the same ability level do not see the same items. NWEA continuously field tests new items each year to expand and maintain the integrity of the item banks. Should minor security breaches or items be exposed, NWEA will remove exposed items from the pool.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.4**

3.4

Describe the procedures used to validate the measures of growth.

The procedures to validate MAP as a growth measure have taken several forms. Formal investigations have been conducted into the construct validity and measurement invariance of MAP tests (Wang, McCall, Jiao & Harris, 2012<sup>1</sup>). In addition, the construct of longitudinal achievement using a vertical scale (RIT reading and mathematics scale, specifically) was validated using a latent growth modeling approach (Wang, Jiao & Zhang, 2012<sup>2</sup>). These studies provide strong support for the validity of the RIT scales as a platform to measure achievement growth. More traditional forms of validity such as concurrent validity have been difficult to initiate, primarily because educational entities commonly test students routinely only with a state sponsored accountability test. Such tests often assess at very gross time periods, making concurrent validity studies subject to key limitations.

Given that NWEA uses differences in (status) achievement estimates as estimates of achievement growth, it is imperative that status estimates (RIT scores) contain minimal measurement error. SEM serves as a form of first level check on the confidence that can be placed in the score as a valid indicator of the student's ability. However, NWEA also monitors other indicators that contribute to making such judgments. Considerable attention is given to Individual Score Validity (ISV) by analyzing student effort demonstrated throughout the test event (Kingsbury & Wise, 2012<sup>3</sup>; Wise, 2012<sup>4</sup>; Wise, Kingsbury & Hauser, 2012<sup>5</sup>; Wise, Ma & Theaker, 2012<sup>6</sup>). Effort is operationalized using a set of heuristic flags triggered by combinations of rapid response times and incorrect answers in different portions of the test. Planning is currently underway for incorporating these ISV flags into production test scripts and reports. By helping to ensure test score validity, arguments for the validity of inferences that depend on growth are strengthened.

In addition to the measures NWEA takes to ensure score validity, schools, and school divisions are encouraged to undertake a number of procedures aimed at enhancing validity during test administration. The primary threats to validity that these procedures address include:

- Attempt to view the secured item pool prior to testing or to record items as testing occurs. The most common way this threat is realized is by teachers or administrators enrolling pseudo-students so that educators can take the test.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.4**

- Retesting students solely to improve a growth score. Since the primary purpose of interim testing is to provide information to help the teacher plan instruction, we believe very strongly that schools should not retest students with valid scores. Threats to validity emerge when retesting procedures are not consistently followed (for example if educators “cherry-pick” retesting to try and inflate their growth scores).
- Inappropriate help or accommodations during the test. Providing help to students that is not specifically authorized in an accommodation.
- Students who are not assigned to teachers within the accountability system.

We recommend that the following safeguards be in place in order to assure the integrity of the testing process, particularly when results are to be used for teacher evaluation. Most of these procedures are not unique to the MAP assessment; they are procedures we would recommend for administration of any test.

- Authorize only one person, preferably a non-teacher, in each school to enroll students (i.e., add student profile data) in the testing system. Most enrollment is done through the central office, but schools need the ability to enter new students into the system, sometimes at the last minute.
- Adopt a policy that prohibits enrolling non-students in the MAP system for the purpose of enabling teachers, parents, or others to take practice tests (thus exposing the item pool) and require that all students have a unique division- or state-issued ID. The division should periodically audit official enrollment records against the record of tested students to confirm that this policy is being followed.
- The MAP system provides data to schools about the validity of individual test events that can be used to ensure the measurement is accurate for each student. These data include test duration, percent correct (on adaptive tests, nearly all students should get approximately half the items right), and when required item-response data. We encourage schools to retest students when these data indicate that an administration is likely to have produced an invalid result. Adopt a policy to govern retesting of students in the MAP system, set specific empirical criteria for testing (for example, all students with test durations under fifteen minutes or percent correct under thirty-five percent should be retested), and require that policy be followed each term. Retesting requests should be approved by the principal, and the division should randomly audit records to assure that retesting policies are being followed.
- Require that accommodations be limited to students with Individual Education Plans and that accommodations allowed be limited to those specifically identified in the plan. Accommodations offered should be noted and approved by the principal for each student.
- Adopt a policy that prohibits teachers from transcribing or saving items from the test and should communicate that policy to teachers.
- MAP testing is normally conducted in a computer lab. We recommend that the test be proctored by both the teacher and a second person (normally the educator or assistant manning the lab). This discourages teachers offering inappropriate help to students. It also has several other benefits. First, the presence of the person responsible for the lab assures that someone in the room is available to troubleshoot problems if they occur. Second, it assures that the teacher has an opportunity to monitor testing conditions to ensure they are appropriate and that the test is properly administered.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.4**

- If a computer lab is used for testing, we also encourage recording and preserving test activity on unobtrusive security cameras. This policy helps the division better assure test security, and also provides the testing procedures integrity in the event that a teacher is questioned on this. The division should establish procedures to make sure assignments of students to teachers are accurately recorded in order to ensure the integrity of results. These procedures should require that all students are accounted for, and should require validation of teacher assignments by both the teacher and the school principal.

**Section 4: Test Administration Procedures**

**Requirement: Section 4: Test Administration Procedures 4.1**

4.1

Describe the administration procedures necessary to produce growth scores. For example, is the assessment designed to be administered multiple times during the year or administered once in the fall and once in the spring?

MAP for Science assessments are a measure of both status and growth. Beginning with the 2011 Norms Study, a student's achievement and growth can be referenced from an estimated growth trajectory the student shares with his/her academic peers (i.e., those students who performed equally well when tested during the same instructional week). Given the student's observed RIT score (and its accompanying SEM) from a test occurring at the specific instructional week, the student's normative performance on the test may be directly determined. The new approach may also be used to predict the student's performance if he/she had tested a week earlier, or several weeks later. As a result, performance may be evaluated at alternative and various points on the relevant instructional calendar and, in this sense, is less rigid than the more common test term window designations would suggest. The new procedure facilitates the determination of the Conditional Growth Index (CGI) which enables the user to directly determine the growth by the student's second testing occasion relative to the student's previous performance, and when compared with academic peers.

The depth of MAP for Science assessment item pools for Virginia is sufficient for three test administrations per year. Our partners typically administer MAP for Science tests during the fall, winter, and spring test terms. NWEA provides four default testing window dates for the purpose of establishing data norms; however, Virginia LEAs will have the flexibility to select test term windows within the following default windows:

- Fall: August 15th through November 30th
- Winter: December 1st through February 28th
- Spring: March 1st through June 15th
- Summer: June 16th through August 14th

**Requirement: Section 4: Test Administration Procedures 4.1**

While flexible, it is important to establish beginning and ending dates for MAP testing for the purposes of data accuracy and division-level reports. The period of time defined by these dates is known as a test term window. The test term window is useful for setting an end date at which an LEA will declare the testing term complete and subsequently have the option to order division-level summary reports. NWEA reports will only display student results that occurred within a division's defined test term window.

NWEA offers Virginia divisions the following recommendations when selecting test window dates:

- Strive for a three week or less testing window to provide comparisons based on students that have received about the same amount of instruction.
- Keep test window dates consistent from one academic year to the next to allow for valid growth comparison from year to year.
- Establish test window dates using the following recommendations to utilize normative data populated into reports:

Time Category	NWEA Recommendation
Fall test window timing	Beginning of academic year (weeks 1-7)
Winter test window timing	Middle of academic year (weeks 15-21)
Spring test window timing	End of academic year (weeks 28-34)
Time between fall and spring testing	Roughly 32 weeks of instruction
Time between testing in consecutive fall or spring terms	Roughly 36 weeks of instruction

**Requirement: Section 4: Test Administration Procedures 4.2**

4.2

Describe any processes used for pre-identifying and/or registering students for testing. Include what data, including the State Testing Identifier, are collected for each student, how data are collected or transmitted, and how data are maintained and securely managed.

NWEA acknowledges the importance of rostering students prior to testing to ensure the right student receives the right assessment. As a result, NWEA offers comprehensive rostering processes and systems to address partner needs. To ease test administration, NWEA's rostering processes and systems are implemented to support finalizing registration in advance of test sessions. For those cases where a student must be registered immediately prior to testing, both MAP platforms provide a solution to roster individual students. For processing larger roster, or program files, NWEA proposes the following process and systems solution.

## Requirement: Section 4: Test Administration Procedures 4.2

### Web-Based Platform

Using the Administration and Reporting Center interface, division users will have control over the process of importing division-wide or school-wide data files to the web-based platform with student, teacher, class, school, and program information, including the State Testing Identifier. Division users assigned the Data Administrator role can upload complete or partial roster files and programs files to the system. Accommodation needs for individual students can be included in the roster files and will be tracked by the system.

The data files uploaded through the Administration and Reporting Center must adhere to the NWEA roster file and programs file templates; these templates define data fields, character limitations, and other restrictions based on business rules that enable the system to add and update existing information in the system with imported data. The roster files and program files templates are provided in Appendix M. Roster files and program files must be in the CSV (comma-separated values) file format to be imported into the web-based assessment system and can be accepted from any SIS system that allows for configurable data exports.

NWEA's web-based assessment system provides end users the flexibility to import all roster data at one time, or to submit partial data imports. The Assessment and Reporting Center interface accepts the following types of roster imports:

- **Instructor-Only Import** allows instructor access to tutorials, the student warm-up, user guides, and instructional resources, such as DesCartes: A Continuum of Learning, before classes begin. This is especially helpful for new instructor orientation. Divisions might also use this import type to update instructor names and e-mail addresses as needed.
- **Student-Only Import** is convenient for schools that need to prepare students for testing within a very short timeframe when instructor and class data is not yet available. This data can be augmented by updating individual student profiles or importing an updated roster file with more complete data. Note that students imported with minimum information required for testing (student name and grade) will not appear on reports until their profiles include all reporting attributes.
- **Complete Import** prepares student data so that students are ready for testing, organizes data for reporting, and provides instructor access to their students' reports.

Imported files undergo multiple data validation processes. If the system detects errors in the file format (for example, the data file is not in CSV format) or inconsistencies in the data (for example, the same Student ID is assigned to two different students), the system generates an error report. The data administrator corrects the errors in the source system – typically the division's Student Information System (SIS) – and submits an updated file. Updated roster files and programs files can be submitted via the Administration and Reporting Center interface as often as is needed.

### Client-Server Platform

On the client-server platform, the student enrollment process is similar to that of the web-based platform described above, except:

**Requirement: Section 4: Test Administration Procedures 4.2**

- Instructor management is not supported; all user accounts are managed via NTE Admin
- The roster files, also in CSV file format, have slightly different data fields requirements
- The roster files are submitted via the NWEA Reports Site
- File validation is manually performed by the NWEA Technical Support team

**Data Transfer and Storage Security**

NWEA systems' security architecture is based on the underlying security pillars of information security: Confidentiality, Integrity, and Availability. Confidentiality ensures that all NWEA system users can access only data that their permission level entitles them to access. Integrity ensures that unauthorized users are not allowed to tamper or alter any data they do not have access to. Availability ensures that NWEA systems and data are available to authorized users when they need them. The security architecture of NWEA systems is built to provide high security by employing security best principles.

NWEA systems transmit and receive data over the Internet using industry standard 256-bit secure sockets layer (ssl) using a 2048-bit certificate. This approach to secure data transfers is in compliance with the Family Education Rights and Privacy Act ("FERPA") and Children's Online Privacy Protection Act ("COPPA").

**Requirement: Section 4: Test Administration Procedures 4.3**

4.3

Describe all materials needed for test administration and how school divisions will order and obtain sufficient quantities. Include details of test booklets and answer documents for paper/pencil testing (if applicable), test administration manuals, etc. If applicable, identify any test administration materials school divisions would be responsible for supplying locally (manipulatives, copies of test materials, etc).

MAP assessments are administered on a computer, typically in a school computer lab, therefore no printed test or administration materials are required. See Section 4.4.5 for the MAP system technical requirements. See Section 4.4.4 for a description of the resources NWEA provides MAP system users to help understand the administration of the assessments.

**Requirement: Section 4: Test Administration Procedures 4.4**

4.4

Provide examples of the test administration manuals to be used with the assessment(s).

**Requirement: Section 4: Test Administration Procedures 4.4**

NWEA provides multiple resources to help MAP users understand the general administration protocol that applies to all MAP assessments and test delivery platforms (client-server and web-based). These resources extend to general information about how MAP tests function and specific “how to” instructions for administering tests on each test delivery platform. Examples include:

- Guidelines for Teacher and Proctor Duties on Test Days, available on NWEA’s Partner Support website at <http://www.nwea.org/support/article/1042/teacher-and-proctor-duties-testing-days> and included with this application as Appendix N.
- MAP Administration Characteristics of Tests, included in the NWEA Technical Manual and provided as Appendix O.
- MAP Proctor Handbook is available at <http://www.nwea.org/support/article/1150/proctor-handbook-map%C2%AE>.

Additional MAP administration documentation is available on the NWEA Partner Support website at <http://www.nwea.org/support/article-16>.

**Requirement: Section 4: Test Administration Procedures 4.5**

4.5

Describe all technology requirements related to school personnel managing the administration of tests and to students completing tests if assessments include technology-based delivery. Include the minimum and recommended hardware and software requirements and network requirements for test administration by school personnel and test delivery to students. Include how assessments are hosted (e.g., locally, vendor, 3<sup>rd</sup> party). Provide examples of user interfaces for test administration by school personnel and test delivery to students. Include descriptions or examples of test navigation and any test tools (e.g., calculator, ruler, highlighter) available to students for testing.

NWEA is one of the first providers of computer-based testing in the United States. Its long history in assessment technology puts it at the forefront of test design and test delivery systems. With this experience, NWEA can deliver to Virginia divisions a high-quality and reliable test with either a web-based or client-server platform. NWEA recognizes not all of Virginia’s schools are equipped with the Internet access required to use a web-based product, so NWEA also offers our client-server platform. These platforms have successfully administered over 262,800 tests in the Commonwealth of Virginia during the 2011-2012 school year.

**NWEA’s Web-Based Platform**

The web-based platform is an end-to-end online testing solution with components that comply with the highest standards for data security, transfer speed, and disaster recovery, while providing outstanding up-time during scheduled system availability. The web-based platform is accessible to any computer inside or outside of the division that meets the hardware and software requirements provided in Appendix P. Key components of NWEA’s web-based system are:

- **Online Administration and Reporting Center**, providing authorized users role-based access to administration functionality such as: user, student, organization, program, and test

#### Requirement: Section 4: Test Administration Procedures 4.5

data management. Users will also access the Administration and Reporting Center to create and manage student test sessions, access reports, and instructional resources.

- **The Student Testing Center**, a web-based test taking environment where students log in using a standard web browser to take tests that have been scheduled by proctors.
- **Computer-adaptive tests** designed to measure achievement of students as they grow over time. The tests commonly consist of more than forty questions so the system can provide sub-scores that represent the content goals and an overall score for each student.

NWEA plans to add a new capability to our web-based platform to allow for lower grade Internet connections, expected to be available by July 2014. Until the new capability is available, NWEA recommends divisions and schools with lower-grade Internet connections use the client-server platform as an interim solution. NWEA will carefully monitor the progress of high bandwidth internet service as this capacity grows statewide. We will work with the Virginia DOE and the divisions to determine the optimal process to move divisions to our web-based solution.

#### NWEA's Client-Server Platform

The client-server platform is an end-to-end testing solution for administering tests on the TestTaker Client workstation without Internet connection. Four main components comprise the client-server assessment system:

- **Network Test Environment (NTE) folder(s)**. The MAP Assessments are designed to be administered in a local network environment. The NTE is a folder located on the local network that contains agency data, student data, and test package data.
- **NTE Administration Tool (NTE Admin)**. NTE Admin is an installed application used for managing a division's NTE. This program can be used to perform tasks such as adding students to an NTE on a student-by-student basis, moving students from one NTE to another or uploading test results. To use the NTE Admin, you will need an Internet connection and Microsoft .NET Framework 4.0. Since NTE Admin is only used for administrative functions and not for the actual test administration, it is not necessary to have NTE Admin installed on every computer used for test administration. NWEA recommends at least one computer in each school have NTE Admin installed, and that these computers have a high-speed Internet connection.
- **TestTaker Client**. TestTaker is the software application that students use to take NWEA assessments. Proctors will log in to TestTaker to select the test to be administered and the student to whom the test will be administered. Proctors also have access to the Proctor Administration Menu, where they can pause or terminate a test if needed. Running TestTaker requires network access to the NTE server but does not require an Internet connection.
- **NWEA Reports Site**. The NWEA Reports Site is a secure website where division assessment coordinators, school administrators, and teachers can log in to see their students' test results. The Reports Site also includes administrative tools to help the division assessment coordinator manage the testing season, including a secure upload page for Class Roster Files, a page for declaring division testing complete, and a report order page for school and division summary reports.

**Requirement: Section 4: Test Administration Procedures 4.5**

The client-server user interface for test administration and delivery by school personnel is nearly parallel to that of the web-based platform.

Both the web-based and client-server assessment systems are deployed and operated independently for each Virginia division, requiring no state-level mediation. Technical requirements for both MAP platforms are included as Appendix P.

**Requirement: Section 4: Test Administration Procedures 4.6**

4.6

Describe accommodations available to students with disabilities and limited English proficient students. Include procedures related to the provision of accommodations to eligible students.

The adaptive nature of MAP assessments make them appropriate for many students in particular subgroups (i.e., students with disabilities, limited English proficient students, etc.). We recognize that student needs and skills within these groups are not homogeneous. In view of this, we focus accommodations on skills and practices used consistently during instruction and assessment. NWEA-approved accommodations pursuant to state and/or division policy are appropriate for an individual student. The following guidance is available when administering MAP to subgroups requiring accommodations:

**General Guidelines for Accommodations**

- Responsibilities: Scribes, page-turners, educational assistants, and other people supporting a student's test must be neutral in responding to the student during test administration. Assistance in test administration must not lead a student to the correct answer. The student's response must accurately represent the student's own choice.
- Reading Tests: Do not read aloud any portion of the reading test to any student.
- Symbols: Do not pronounce or explain math or science symbols.
- Definitions: Do not define any words within test questions.
- Interpreting Results: While many accommodations are recognized, NWEA has not performed studies to confirm how their use affects assessment scores. By definition, accommodations should allow more precise and more valid estimates of true scores. That is, accommodations are intended to "level the playing field" for students with known disabilities. They should not advantage any student relative to any other student at the same true ability level. However, for certain presentation accommodations noted below (2 and 3), we cannot state with a high level of certainty how validity of test scores is affected due to the potential change in the construct being measured. This qualification should be considered when interpreting test scores or using them to make important educational decisions.

**Presentation Accommodations**

1. Simplify or clarify directions
2. Native language translation (oral or signing) of test directions  
NOTE: See "Interpreting Results" under the general guidelines above.

**Requirement: Section 4: Test Administration Procedures 4.6**

3. Native language translation (oral or signing) of test questions (not answer options) for math, science, or language usage tests  
NOTE: See “Interpreting Results” under the general guidelines above.
4. Read or reread aloud the test directions
5. Read or reread aloud the test questions (not answer options) for math, science, or language usage tests only  
NOTE: See “Symbols” under the general guidelines above.
6. Use visual magnification devices or software (for example, MAP is compatible with ZoomText or MAGic)
7. Use auditory amplification devices, noise buffers, or software
8. Use masks to block portion of screen; for example, the student may use a sticky note, index card, or a blank sheet of paper to move down the screen as he or she is reading

**Response Accommodations**

1. Assign scribe to record responses
2. Dictate responses to a scribe
3. Point to responses for a scribe
4. Respond in native language

**Setting Accommodations**

1. Test an individual student in a separate setting
2. Test a small group of students in a separate setting; for example, in a Title I room or counselor’s office
3. Minimize distractions; for example, use a study carrel

**Timing/Schedule Accommodations**

1. Administer test over multiple sessions in a day
2. Administer test over a number of days
3. Administer test at a particular time of day
4. Allow flexible schedule
5. Extend time allotted by proctor or test administrator (All MAP assessments are designed to be untimed.)
6. Allow stop-the-clock breaks
7. Administer at time of day most beneficial to student
8. Offer breaks

**Materials or Devices Accommodations**

1. Provide scratch paper

**Requirement: Section 4: Test Administration Procedures 4.6**

2. Availability of a calculator (when appropriate, it is provided on the screen)

**Miscellaneous Accommodations**

1. Provide drink during testing
2. Provide snack during testing

In addition to these accommodations, MAP is compatible with third party software to provide capabilities to address the needs of visually impaired students including magnification, highlighting, and color contrast. NWEA will explore MAP compatibility with third-party software to accommodate the needs of blind students by evaluating third party software compatibility of text to speech and electronic Braille.

The following table identifies which English Language Learner (ELL) accommodations to the MAP assessment are available within the assessment:

- Read directions orally in English
- Allow flexible schedule – test in sessions, over several days
- Extended time
- Provide scratch paper/allow to mark test booklet
- Administer in a small group
- Administer individually

**Requirement: Section 4: Test Administration Procedures 4.7**

4.7

Describe procedures for completed student tests to be submitted for scoring and reporting purposes.

Since 2000, all computer-administered NWEA assessments have been computer-scored by the MAP assessment system as tests are administered. Computer scoring allows NWEA to provide preliminary testing results for the student and teacher at the end of each test to identify students needing intervention in a timely manner and to accurately point to instructional learning objects.

Scoring of MAP assessments is based on Item Response Theory (IRT). Student scores are calculated, not only by employing an algorithm that each student is administered. NWEA's system has been rigorously tested to ensure scores rendered from the system are calculated and displayed correctly.

## Section 5 : Scoring and Reporting

### Requirement: Section 5: Scoring and Reporting 5.1

#### 5.1

Describe scoring procedures for all item types and test forms administered, including implemented quality control measures.

MAP assessments are adaptive and employ a common scoring algorithm. During the assessment, a Bayesian scoring algorithm is used to inform item selection. Bayesian scoring for item selection prevents the artificially dramatic fluctuations in student achievement at the beginning of the test that can occur with other scoring algorithms.

Although the Bayesian scoring works well as a procedure for selecting items during test administration, Bayesian scores are not appropriate for the calculation of final student achievement scores. This is because Bayesian scoring uses information other than the student's responses to questions, such as past performance, to calculate the achievement estimate. Since only the student's performance on the day he/she takes a test should be used to give the student's current score, a maximum-likelihood algorithm is used to calculate a student's actual score at the completion of the test.

NWEA uses the Rasch Item Response Theory model to create its vertical scales, called RIT scales. MAP results, reported as RIT scores, relate directly to the curriculum scale in each subject area.

At inception, NWEA adopted Item Response Theory in general and the Rasch model in particular to create its highly stable vertical RIT scales. There is one RIT scale each for mathematics, reading, and language usage. Using the RIT scale to report test results makes it possible to follow a student's educational growth from testing season to testing season and year to year. Please note that each subject area has a unique alignment to the RIT scale; as a result, scores between subjects are not equivalent.

By using Item Response Theory to create the scales and anchoring item difficulty estimates to them, the RIT scales are comparable from one set of items to another, and from one set of examinees to another. This enables comparisons of the scores from different students, or from the same student at different times, even though different sets of test items are administered. This also allows longitudinal comparison of student performance to be made.

#### **QA Process**

NWEA's dedicated QA staff work together with NWEA's research team to perform a comprehensive review of all reporting data for accuracy and quality. This includes the validation of all business rules and formulas applied when generating results reports for both standard reports provided via the assessment platform and all custom reports or data extracts provided.

**Requirement: Section 5: Scoring and Reporting 5.2**

5.2

Describe the type of reporting provided (e.g, static and/or dynamic, electronic and/or paper-based, item-level, strand-level, and/or test-level scoring). Include approximate timelines for score reports to be available to divisions, how score reports will be accessed and/or obtained, and samples of student, class, school, and division score reports and sample record layouts for electronic data files.

NWEA assessments provide a comprehensive set of computer-based reports, instructional resources, data tools and comparative information sources to gain insight into student growth and proficiency at the individual, classroom, school, grade and division level. System reports and instructional resources are student-centric, research-based, and data-driven. They provide data to inform instruction, evaluate programs, justify budget initiatives, and make key decisions in the classroom, school, and division levels.

These analytical tools include the following reports:

- Student-level reports including the Student Progress Report and the Student Goal-Setting Worksheet provide current and historical student performance. This includes growth and growth projections as well as normative data and Lexile scores to assist students, teachers, and parents select suitable reading texts. Student Progress Reports can be printed by the educator and shared with parents to create a parent/teacher dialog around their student's learning, performance, and growth. To help guide parents through the various components of this report, a special one page annotated version of the report is provided as an accompaniment. The Student Goal Setting Worksheet is especially helpful in allowing parents to see the growth potential of their child.
- Class and teacher reports provide an overview of class performance and detailed information about each student, including: assessment and goal area performance, norm percentile rank, growth, and predictive performance on the summative assessments once a scale alignment study is completed. Teachers can use these reports to differentiate instruction and identify areas of strength and weakness at a group level.
- School-level reports provide information that can help administrators evaluate programs and foster collaboration.
- Grade-level reports provide summary information to help school administrators assess trends, identify areas of strength and weakness, and percentage of students meeting their growth targets
- Division-level reports which help partners assess performance trends by grade and school. These reports present longitudinal data for a particular school and are valuable resources in planning and monitoring school improvement plans.

NWEA uses a user-centric approach to report design. Throughout the product lifecycle, usability testing is incorporated to understand how a teacher or administrator intends to use the reports, and incorporates this mindset in the development of report designs that are further enforced by the experience of NWEA's research teams and education experts. This user-centric approach provides greater benefit to partners through limiting the amount of time required by the user to learn the content of the reports. The information contained within our reports is intended to aid in

**Requirement: Section 5: Scoring and Reporting 5.2**

swift and decisive action plans. The MAP assessment system scores each test as it is administered and provides preliminary results for the student and teacher at the end of each test. Reports are generated and accessed online.

Longitudinal Results: NWEA collects and stores student MAP assessment results across time, providing educators with longitudinal data to compare student growth over time. Because MAP results are reported on vertical scales called the RIT scales, it is possible to compare individual students’ results over time and compare sets of results. Furthermore, educators can access reports for the current academic year and the previous academic year from NWEA’s Reports Site (see page 5 of Appendix Q for a sample ).

DesCartes: A Continuum of Learning: For students in grades 2-12, NWEA offers DesCartes: A Continuum of Learning. DesCartes translates student goal scores into relevant skills and concepts by achievement level and are aligned to Virginia’s Goal Structures created from the Virginia SOL. DesCartes can be used to form and implement flexible grouping strategies, and identify and monitor student needs according to student progress.

**Availability and Format of Reports**

The MAP assessment system scores each test as it is administered and provides preliminary results for the student and teacher at the end of each test. Reports showing test results are available to teachers and administrations via login at the web-based reporting interface. For the Web-based MAP system, reports are accessed via the MAP Administration and Reporting Center. Please refer to the table below for complete information regarding reports availability. Reports can either be viewed online and/or using a PDF viewer, such as Adobe Reader®. From the PDF viewer reports can be viewed, saved, or printed.

Available Immediately	Available within 24 hours*
<ul style="list-style-type: none"> <li>▪ Class Breakdown by RIT Report</li> <li>▪ Class Breakdown by Goal Report</li> <li>▪ Class by Projected Proficiency</li> <li>▪ Descartes: A Continuum of Learning</li> <li>▪ Primary Grades Instructional Data</li> <li>▪ Student Goal Setting Worksheet</li> </ul>	<ul style="list-style-type: none"> <li>▪ District (Division) Summary Report</li> <li>▪ Student Progress Report</li> <li>▪ Class Report</li> <li>▪ Grade Report</li> <li>▪ Achievement Status and Growth Reports</li> <li>▪ Student Growth Summary Report</li> <li>▪ Projected Proficiency Summary Report</li> </ul>

\*These reports may be available immediately depending on the amount of data.

In the client-server MAP system, reports are accessed via the NWEA Reports Site. Please refer to the table below for information about reports availability timeframe and format.

**Requirement: Section 5: Scoring and Reporting 5.2**

Available Reports	Timeframe	Format
Teacher- and class-level reports	The day after test results are uploaded to NWEA	PDF*
Grade-, school-, and district-level (division-level) reports	Within 24-72 hours after requesting reports	Crystal Reports*, tab-delimited (Data Text File), and CSV (Comprehensive Data Files)
Dynamic Reports showing class-, school-, grade-, and district-level (division-level) growth and proficiency data	One week after test results are uploaded to NWEA	Online display; can be exported to PDF

**Report Samples**

Samples of student, class, school, and division reports are included in Appendix Q, Annotated Reports. In addition to sample reports, Appendix Q provides guidance on how to interpret these reports. The annotated reports display sample versions along with helpful explanations of what each of the different report components mean. NWEA partners use the Annotated Reports to gain better understanding for how to interpret assessment scores. Additionally, NWEA offers onsite workshops and online training specifically designed to help educators interpret results and adjust instructional practices to meet their students' academic needs. See Proposal Narrative, Section a. for information about NWEA's online support and professional development workshops.

**Requirement: Section 5: Scoring and Reporting 5.3**

5.3

Describe all data tools available to school division staff for the analysis of data and the creation of customized reports.

NWEA is committed to providing accurate and timely results to Virginia educators. In addition to offering the aforementioned menu of reports, NWEA also makes available two additional options for obtaining data: raw data files and custom reporting.

**Raw Data Files**

For divisions, the MAP system generates a Comprehensive Data File that contains assessment results, student data, class assignments, and special program (sub-group) designations in a flat text format (common-separated values (CSV)) for further analysis and use in external data or reporting systems, i.e. data warehouses. See Appendix M for file format sample.

**Optional Division Custom Reporting**

If Virginia divisions require custom reporting capabilities, NWEA is able to deliver sophisticated custom reports designed alongside our partners at a negotiated rate.

- 
- <sup>1</sup> Wang, S., McCall, M., Jiao, H., & Harris, G. (April, 2012). Construct Validity and Measurement Invariance of Computerized Adaptive Testing: Application to Measures of Academic Progress (MAP) Using Confirmatory Factor Analysis. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.
- <sup>2</sup> Wang, S., Jiao, H., & Zhang, L. (2012, April). Validation of Longitudinal Achievement Constructs of Vertically Scaled Computerized Adaptive Tests: A Multiple Indicator Latent Growth Modeling Approach. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.
- <sup>3</sup> Kingsbury, G. G., & Wise, S. L. (in press). Turning the page: How smarter testing, vertical scales, and understanding of student engagement may improve our tests. In R. W. Lissitz and H. Jiao (Eds.) *Computers and their impact on state assessment: Recent history and predictions for the future*. Charlotte, NC: Information Age Publishing, Inc.
- <sup>4</sup> Wise, S. L. (2012). *The utility of adaptive testing in addressing the problem of unmotivated examinees*. Manuscript submitted for publication.
- <sup>5</sup> Wise, S. L., Kingsbury, G. G., & Hauser, C. (2012). *How do I know that this score is valid? The case for assessing individual score validity*. Manuscript submitted for publication.
- <sup>6</sup> Wise, S. L., Ma, L., & Theaker, R. A. (2012, May). *Identifying non-effortful student behavior on adaptive tests: Implications for test fraud detection*. Paper presented at the Conference on the Statistical Detection of Potential Test Fraud, Lawrence, Kansas.



**Offeror Name:** Northwest Evaluation Association (NWEA)

**Proposed Assessment Name:** MAP for Primary Grades Assessments

**Content Area(s) and Grade Level(s) Assessed:** Mathematics and Reading for Grades K-2

## Section 1: Overview of Tests

### Requirement: Section 1: Overview of Tests 1.1

#### 1.1

Describe the specific grade(s) and subject area(s) covered by each assessment and provide an overview of the content and skills measured. Include the types of test items used, the mode(s) of delivery, the availability of equivalent forms, including short forms or screeners (if available) and a test blueprint for each test being proposed.

MAP for Primary Grades is a computer administered assessment appropriate for administration at grades K-2. MAP for Primary Grades tests contain selected-response items that are interactive in nature, meaning students can manipulate and construct answers based on the learning being assessed. Instead of selecting from a list of answers, the student is often able to use a mouse to perform an action or construct a response. MAP for Primary Grades is a useful tool for screening and placing students as well as monitoring growth.

Three test types are provided with MAP for Primary Grades. Multiple assessments enable teachers to administer tests appropriate to each student's readiness. The different MAP for Primary Grades test types are described below:

- **Screening Assessments:** The Early Numeracy and Early Literacy Screening assessments are designed as an initial screening for students as they enter school (typically in kindergarten or first grade). This initial assessment helps to provide immediate screening regarding basic numeracy and literacy concepts and skills. The screening test is designed to adjust to more challenging or more basic questions depending on the need of the student as he or she proceeds through the assessment. Reports for the Screening assessment display the percent correct for each element of the identified content.
- **Skills Checklist Assessments** are used to inform instruction relative to basic reading and mathematics skills. The length of the Skills Checklist assessments adjust based on student responses to test items. Students who take a reading Skills Checklist test are randomly presented with all items in the assessment. In mathematics, the Skills Checklist tests randomly present items and automatically stop after the first ten items if the student has not scored at least sixty percent by that point in the test. Students who answer sixty percent or more of the first ten items correctly will then see all the remaining items in the test. Skills Checklist tests report number correct and can be administered as often as is useful to the educator.
- **Survey with Goals Assessments:** The MAP for Primary Grades Survey with Goals adaptive tests are designed to provide educators with instructional information about what students are ready to learn and to measure academic growth in reading and mathematics. These tests, which are appropriate for students who have a firm grasp of foundational skills, adapt to the

**Requirement: Section 1: Overview of Tests 1.1**

level of difficulty for each student. Survey with Goals tests provide an overall RIT score for the subject as well as sub-scores in each of the goal performance areas. The overall subject score is used to compute student growth using NWEA's RIT scale. The MAP for Primary Grades Survey with Goals tests can be administered up to three times per year.

The MAP for Primary Grades assessments for grades K-2 are created from large item pools. The MAP for Primary Grades assessments meet the unique needs of early learners by displaying interactive elements and providing audio for beginning readers. Audio is used through the MAP for Primary Grades tests to help beginning readers understand the tasks presented in the assessment questions.

By spring 2013, NWEA's Virginia State-aligned MAP for Primary Grades for Mathematics and Reading assessments will be available for administration to students in grades K-2. The new MAP for Primary Grades will be aligned to the Virginia Standards of Learning. See Appendix X for a sample MAP for Primary Grades Test Blueprint. Upon completion of the new Virginia State-aligned MAP for Primary Grades assessments, NWEA will provide the Virginia DOE with new test blueprints.

**Requirement: Section 1: Overview of Tests 1.2**

1.2

Provide evidence of alignment of test items to the Virginia Standards of Learning (SOL) for existing assessments. For assessments developed in response to the RFP, provide a plan for assuring the alignment of test items to the SOL.

For the creation of computer adaptive tests tightly aligned to the Virginia Standards of Learning (Virginia SOL), NWEA will conduct an item-by-item analysis of all items in the reading and mathematics item banks to manually align items to the standards. As a result of this work, each item in the item pool for the Virginia State-aligned MAP for Primary Grades tests will have a confirmed alignment to the Virginia SOL. Test content alignment is conducted by NWEA Content Specialists. These Content Specialists are content area experts with five or more years of classroom teaching experience. They have an expert level of knowledge of writing and reviewing items for classroom, district, interim, and state assessments; have an expert knowledge of national trends and initiatives in assessment. Additionally, they have experience creating test structures aligned to state standards and selecting item pools for those assessments.

NWEA's alignment process is the cornerstone of NWEA's ability to deliver computer-adaptive, grade-independent tests that cover a wide range of content and have sufficient depth of items to provide data on a wide range of student performance at any given grade level.

The process that will be used to develop MAP for Primary Grades tests aligned to the Virginia SOL includes six steps:

**1. Review of the Virginia Standards of Learning (Virginia SOL).** Content Specialists will analyze the Virginia SOL to develop an overarching understanding of what the state expects students to know and be able to do within a content area.

### Requirement: Section 1: Overview of Tests 1.2

**2. Creation of Cross-Grade Level Goal Structure.** Content Specialists will create a cross-grade level goal structure of the Virginia SOL for each subject area. A goal structure is created by combining the Learning Standards across grade levels to create a two-tier framework consisting of goals and sub-goals. These goal structures are used to create the tests and reports.

**3. Alignment of Learning Continuum Statements to the Goal Structure.** Content Specialists will conduct an alignment of the NWEA Learning Continuum statements, which act as holding places for large groups of items representing specific skills or concepts, to each appropriate grade-level learning standard in the Virginia SOL. All goal structures will receive a final review by at least one subject-area Content Specialist that was not involved in the initial process.

**4. Item-by-Item Alignment.** In order to validate the alignments suggested by the correlation of Learning Continuum statements to the Virginia SOL, Content Specialists will check whether each item is aligned to the most appropriate grade-level standard and makes changes as necessary. This step will assure a quality alignment.

**5. Psychometric Review of Selected Test Pool.** Psychometricians in NWEA's Measurement and Data Services Department will conduct a pool analysis to validate that the depth of the item pools for each test is sufficient to administer the tests multiple times per year, depending on the specific test.

**6. Creation of Tests.** The Test Production team will create test packages so the tests can be administered to Virginia students.

Following the six-step alignment process outlined above, NWEA expects to release a version of MAP for Primary Grades that is aligned to the Virginia SOL in spring 2013. NWEA has significant experience creating state-specific tests. State-specific tests refer only to the items in the larger item pool selected to assess the learning standards for that state. The goal structure employed in assessments and reports mirrors the organization of the state standards so that teachers can readily determine their students' needs within their curriculum. All items in the total item pool are calibrated against the stable and reliable RIT scale, independent of their alignment to specific state standards. Therefore state-aligned tests have technical properties identical to NWEA's nationally-aligned MAP for Primary Grades tests. Specifically, measures of validity, reliability, precision, distribution of scores, and distribution of growth are consistent across variously aligned test versions. State-aligned tests ensure students are assessed on the constructs and skills that have been defined in their local learning standards and teachers are able to interpret data in the context of the standards to which they are accountable.

Because development of the NWEA Virginia State-aligned MAP for Primary Grades has not yet begun, evidence of alignment is not available at this time. NWEA will develop evidence of alignment similar to the evidence submitted with this application for MAP for Mathematics, Reading, and Language Usage (see Appendix G). The new evidence will be made available to the Virginia DOE once test development is complete, which is expected July 2013.

NWEA submits Appendix Y as a representation of the content assessed by the existing item pools of the nationally-aligned MAP for Primary Grades assessments. To develop the nationally-aligned MAP for Primary Grades assessments, NWEA conducted extensive content research related to foundational skills in early numeracy and early literacy development during the creation of MAP for Primary Grades. Skills included in MAP for Primary Grades assessments were chosen based on fundamental skills considered essential in most state standards in

**Requirement: Section 1: Overview of Tests 1.2**

kindergarten through grade two. These item pools allow MAP for Primary Grades tests to provide precise achievement estimates (standard errors typically below 3.5 RIT points) for students in grades K-2. To create the Virginia-aligned MAP for Primary Grades assessments, items will be hand-aligned from these pools and new items will be created, as necessary, to provide adequate coverage of the Virginia standards. This documentation is provided for illustrative purposes only and does not represent a Virginia alignment. Evidence of alignment to Virginia standards will be available with the release of the aligned test.

**Section 2: Technical Characteristics**

**Requirement: Section 2: Technical Characteristics 2.1**

2.1

Provide evidence of content, construct, concurrent, and predictive validity as appropriate. Include validity evidence that supports the use of scores from the proposed assessment in teacher evaluation, addressing specifically the validity of using assessment results to support inferences about effectiveness of teacher in producing growth in student performance (if available).

To determine content validity for MAP for Primary Grades, a content framework was developed based on extensive curricular and instructional research regarding the skills and concepts deemed critical for success in early literacy and numeracy development. Additionally, items with particular emphasis on the developmental needs of young learners were developed. For the purposes of this application, NWEA has included validity evidence for its MAP tests (grades 2-12) as Appendix Z. Appendix Z includes information about content, concurrent, predictive, and criterion-related validity.

**Content Validity**

Skills included in MAP for Primary Grades assessments were chosen based on fundamental skills considered essential in most state standards in grades Pre-K-2. Additionally, the mathematics concepts and skills within MAP for Primary Grades Mathematics assessments, published in 2006, reflect the most recent publication of the NCTM's Focal Points for Grades Pre-K-8 noting that these concepts are "the most important mathematical topics for each grade level. They comprise related ideas, concepts, skills, and procedures that form the foundation for understanding and lasting learning." (2009, NCTM)

**Construct/Criterion Validity**

MAP for Primary Grades Survey with Goals assessments have been correlated with the state aligned item pool structures of the MAP assessments. The MAP assessments are nationally administered assessments that have normative data associated with them. See Appendix Z for these correlational data.

**Requirement: Section 2: Technical Characteristics 2.1**

Because most state accountability tests do not assess students in grades K-2, predictive validity evidence specific to MAP for Primary Grades is not available. However, because MAP for Primary Grades and MAP assessments use the same stable, equal-interval scale, scores on MAP for Primary Grades assessments, when viewed in light of normative expectations, can provide valuable information about students' progress toward proficiency levels as measured by accountability tests that begin in third grade.

**Requirement: Technical Characteristics 2.2**

2.2

Provide evidence of reliability, both for the total test and for any subtests for which scores are reported. Include estimates of error in measurement.

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

Requirement: Technical Characteristics 2.2
[Redacted]

Requirement: Technical Characteristics 2.3
2.3 Provide evidence that the assessment is appropriate for use with student subgroups, including English language learners and student with disabilities. Include documentation that the assessment does not exhibit bias toward any major subgroups (e.g., through an analysis of differential item functioning). In addition, provide a sensitivity review to demonstrate the assessment tasks and items are designed to be accessible and fair for all students.
The adaptive nature of MAP assessments make the tests appropriate for student subgroups with a wide range of skills and needs. When used consistently during instruction and assessment, NEWA approved accommodations pursuant to state and/or division policy are appropriate for an individual student.
<b>Use of MAP with Subgroups</b> The adaptive nature of MAP assessments allows it to serve as an appropriate academic achievement measure for a broad range of students without modifications or accommodations. This includes many, but not all students with disabilities. The MAP system measures students by providing questions that match each student’s current level of challenge. This allows students with mild cognitive difficulties to be measured accurately by the tests.

### **Requirement: Technical Characteristics 2.3**

In addition, a wide variety of accommodations allow the tests to be given in a manner designed to preserve the accuracy of the measurement. A list of allowable accommodations is provided in Section 4.4.6.

Although students with disabilities such as specific learning disabled, mildly delayed, communication disorder, behavioral disorder, and mild motor impairment may take MAP assessments, collection of this level of detail about specific students is provided at the discretion of each school division. LEAs vary considerably in their willingness to provide such information routinely and in a standard format that would make it useful beyond the LEA providing it. For this reason, NWEA has not been in a position to initiate a comprehensive study of validity and reliability of MAP assessments for students with disabilities. Should the Department be interested in participating in such a study, NWEA would be pleased to work with division staff to identify methods and processes for collecting data to conduct an analysis on these special populations of students.

To establish validity and reliability evidence for the proposed MAP tests use for English Language Learners (ELL), several additional forms of data would be required. Since ELL students will vary in their exposure to and instruction in English, it would be important to verify differences in these areas. It is expected that this could be accomplished by accessing extant data sources such as student demographic and instructional records. To establish a reference level or progression of English language performance, scores from a single or defined set of common English language proficiency tests would be important. Ideally, the test(s) would have reliability estimates that meet conventionally accepted standards as well as validity evidence supporting their use in at least reading performance across difference ages and stages of development of English language proficiency. It would be desirable, though not absolutely essential, for the tests to be aligned to the WIDA English Language Proficiency Standard. The availability of a large representative sample of such data would allow: 1) validity evidence for the use of Virginia State-aligned MAP tests with ELL students, and 2) a determination to be made about the conditions and circumstances under which these MAP tests would not be valid.

#### **Differential Item Functioning Analysis**

Since the publication of NWEA's 2011 Technical Manual, its Research team has completed a Differential Item Functioning (DIF) Analysis for MAP for Primary Grades, which is included as Appendix CC.

The MAP for Primary Grades Survey with Goals assessments are designed for students in kindergarten through second grade. These assessments include adaptive Survey with Goals tests in reading and mathematics.

The MAP for Primary Grades system uses the same measurement scales that are used in the MAP system, which allows a direct connection between the fundamental skills assessed in the MAP for Primary Grades system and the learning of a student in later years.

Students enter school with a wide variety of life experiences. Early identification of each student's achievement level provides a strong foundation for teachers to use in establishing an environment for academic success. The assessments in the MAP for Primary Grades system are

### Requirement: Technical Characteristics 2.3

designed to:

- Provide information to guide instruction during the early stages of a student's academic career
- Identify the needs of a wide variety of primary grades students, from struggling to advanced learners
- Use engaging test items, interactive elements, and audio to encourage student participation for more accurate results and to help beginning readers understand the test questions



NWEA takes steps to ensure tests are appropriate for students both prior to test construction and after the test is administered. Tests are comprised of items constructed in accordance with strict item writing guidelines that address both sensitivity and fairness. Sensitivity in this context means an awareness of the different things that can distract a student during assessment. Fairness in this context relates to giving each student equal opportunity to answer the item correctly based solely on their knowledge of the item content. Any sensitivity and fairness issues found in items are eliminated in revision prior to test construction.

Each item is evaluated against a set of criteria. An item is flagged if it:

- Requires prior knowledge other than the skill/concept being assessed
- Has cultural bias
- Has linguistic bias
- Has socio-economic bias
- Has religious bias
- Has geographic bias
- Has color-blind bias
- Has gender bias
- Inappropriately employs idiomatic English
- Offensively stereotypes a group of people
- Mentions body/weight issues
- Has inappropriate or sensitive topics (smoking, death, crime, violence, profanity, sex, etc.)
- Has other bias issues

#### **Bias and Sensitivity Review**

As part of the item development process, NWEA conducts a bias and sensitivity review. Sensitivity in this context means an awareness of the different aspects of an item or context that can distract a student during assessment. Fairness in this context relates to giving each student an equal opportunity to answer the item correctly based solely on his or her knowledge of the item

### Requirement: Technical Characteristics 2.3

content. A well-constructed item serves to activate and focus a student's thought process on the task presented in the item. A successful item is fair to all students. An item should NOT:

- Distract, upset, or confuse in any way.
- Require construct-irrelevant or specialized knowledge.
- Favor students from certain language communities.
- Favor students from certain cultural backgrounds.
- Favor students based on gender.
- Favor students based on social economic issues.
- Employ idiomatic or regional phrases and expressions.
- Stereotype certain groups of people or behaviors.
- Favor students from certain geographic regions.
- Favor students who have no visual impairments.

A hard and fast list of potentially distracting or upsetting material does not exist, but there are topics that are seldom appropriate for K-12 level assessments, such as sexuality, illegal substances, illegal activities, excessive violence, discriminatory descriptions, death, grieving, catastrophes, animal neglect or abuse, loss of family member, or weight and body issues.

The best way to ensure that items are as fair and sensitive as possible is to review them specifically for sensitivity and fairness issues. Sensitivity and fairness reviews are incorporated into the Editorial review, Item Production review, and both sets of Content reviews.

### Requirement: Technical Characteristics 2.4

2.4

Provide evidence that the assessment includes items of varying difficulty to ensure accurate measurement of student achievement across the ability continuum, including the tails of the score distribution.

MAP for Primary Grades Survey with Goals tests provide assessment items varying in difficulty with items appropriate for grades K-2 in mathematics and reading (see Section 1.1.1). NWEA's psychometricians validate the item pools to ensure they are deep enough that a student taking a particular test will not see the same test item in a fourteen month time period, even if taking a test three times per calendar year. The MAP for Primary Grades assessments for grades K-2 are created from large item pools.

For an example of additional evidence of the varied difficulty of items that cover the expected ability distribution as well as sufficient item coverage, refer to Appendix G.

### Section 3: Use of Assessment as a Measure of Growth

#### Requirement: Section 3: Use of Assessment as a Measure of Growth 3.1

3.1

Provide evidence that the scores resulting from the assessment have been used as measures of growth by other local or state education agencies.

In January 2012, MAP for Primary Grades assessments were reviewed and approved by the New York State Education Department's list of approved growth assessments of the New York State Common Core Standards for use in teacher and principal evaluations. The list of approved assessments including MAP tests can be viewed on the following website: <http://usny.nysed.gov/rttt/teachers-leaders/assessments/home.html#assess>.

For additional evidence of use by local or state agencies, please see Section d. of the Proposal Narrative.

#### Requirement: Section 3: Use of Assessment as a Measure of Growth 3.2

3.2

Describe the methodology used to measure growth. For example, does the assessment employ a vertical scale, use a computer-adaptive model to measure growth over time, or employ some other methodology. Does the methodology allow for the longitudinal measure of growth across academic years? What about the measurement of required growth on the proposed assessment to reach proficient on the statewide assessments (the Standards of Learning tests) in a specified amount of time? Include standard setting studies or other analyses conducted to establish measures of growth.

NWEA takes a very straightforward approach to estimating growth. Since the RIT scales are vertical (cross-grade) scales, growth is estimated directly based on difference (or change) between the status scores from two or more test events. These difference scores are qualified using the standard errors of the individual status scores involved in the comparison (specifically,  $\sqrt{SEM_{test1}^2 + SEM_{test2}^2}$ ). In order to take this approach, heavy reliance is placed on score precision and scale stability. Theoretically, the length of the interval between scores is only relevant when growth rate or comparison of growth (change) scores is of interest. Given this, growth can be computed and evaluated across two contiguous terms, across a school year, or longitudinally across multiple school years. To help inform the interpretation of observed growth, NWEA has developed a set of national norms to serve as points of reference. The latest (2011) norms continue a practice of publishing norms every three years, from 1996. The 2011 norms study is included on CD with this proposal.

The 2011 NWEA growth norms were developed using several methodological enhancements designed to improve their utility as a means of anticipating as well as evaluating student academic growth. These enhancements include:

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.2**

- The use of empirical instructional weeks in estimating growth rates, which allowed the number of instructional weeks preceding the first test and varying test 1 – test 2 intervals to be used in normative growth projections
- Use of the first (pretest) score as a covariate in calculating normative growth projections
- Weighting scores by their standard errors in calculating normative growth projections
- Use of an empirically derived state-level School Challenge Index (based on NCES data) as a post-stratification weighting variable for both the growth and status norms.

As a result of these enhancements, student growth can be projected and evaluated by taking into account the student’s initial RIT score, the time in the instructional year the first (e.g, fall) test is administered, the time interval between the two tests, and the standard error of the RIT score(s) involved. This provides a great deal of flexibility in how the RIT scale norms are used to help interpret growth as well as how growth standards could be approached and set. For example, a MAP scale alignment study to the Virginia State SOL tests was recently completed. Like other studies of this type, this study provides estimates of RIT scores associated with each cut score point at each grade level for mathematics and reading. Thus, even though a state does not use a cross-grade (vertical) scale, the alignment study will provide “bridges” from state test results to the (vertical) MAP scales. Using these bridges and the 2011 RIT scale growth norms as references, growth standards within and across grades could be meaningfully informed. Such standard setting could take advantage of the enhancements noted above to provide more precise estimates of growth for the purpose of projection or for evaluation. In addition, as a result of the 2011 RIT scale norms study, there has been renewed interest in refining the Hybrid Success Model (Kingsbury & McCall, 2005). The Hybrid Success Model provides an extension of shorter term growth estimates to those involving distal achievement targets (e.g., from spring of grade four to spring of grade seven). This model estimates the amount of growth required each year to meet the distal achievement target, given the student’s current achievement level, the normative growth rate from that level, and the number of instructional weeks until the target will be assessed.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.3**

3.3

Describe the methodologies used to control item exposure so that the accuracy of students’ scores is not impacted by multiple exposures to the same items.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.3**

Because MAP tests are adaptive, all test takers receive a customized form of the assessment each and every time they take a MAP assessment. NWEA controls the items each student sees by monitoring the specific items exposed to each student. Survey with Goals tests can be administered three times per year without the student seeing the same item until fourteen months have elapsed. In addition, the depth of the item pool makes it unlikely that students will encounter the same item again even after two years. Test algorithms assure items in the same content and difficulty category are chosen randomly, so that the students at the same ability level do not see the same items. NWEA continuously field tests new items each year to expand and maintain the integrity of the item banks. Should minor security breaches or items be exposed, NWEA will remove exposed items from the pool.

**Requirement: Section 3: Use of Assessment as a Measure of Growth 3.4**

3.4

Describe the procedures used to validate the measures of growth.

The procedures to validate MAP as a growth measure have taken several forms. Formal investigations have been conducted into the construct validity and measurement invariance of MAP tests (Wang, McCall, Jiao & Harris, 2012<sup>1</sup>). In addition, the construct of longitudinal achievement using a vertical scale (RIT reading and mathematics scale, specifically) was validated using a latent growth modeling approach (Wang, Jiao & Zhang, 2012<sup>2</sup>). These studies provide strong support for the validity of the RIT scales as a platform to measure achievement growth. More traditional forms of validity such as concurrent validity have been difficult to initiate, primarily because educational entities commonly test students routinely only with a state sponsored accountability test. Such tests often assess at very gross time periods, making concurrent validity studies subject to key limitations.

Given that NWEA uses differences in (status) achievement estimates as estimates of achievement growth, it is imperative that status estimates (RIT scores) contain minimal measurement error. SEM serves as a form of first level check on the confidence that can be placed in the score as a valid indicator of the student's ability. However, NWEA also monitors other indicators that contribute to making such judgments. Considerable attention is given to Individual Score Validity (ISV) by analyzing student effort demonstrated throughout the test event (Kingsbury & Wise, 2012<sup>3</sup>; Wise, 2012<sup>4</sup>; Wise, Kingsbury & Hauser, 2012<sup>5</sup>; Wise, Ma & Theaker, 2012<sup>6</sup>). Effort is operationalized using a set of heuristic flags triggered by combinations of rapid response times and incorrect answers in different portions of the test. Planning is currently underway for incorporating these ISV flags into production test scripts and reports. By helping to ensure test score validity, arguments for the validity of inferences that depend on growth are strengthened. In addition to the measures NWEA takes to ensure score validity, schools, and school divisions are encouraged to undertake a number of procedures aimed at enhancing validity during test administration. The primary threats to validity that these procedures address include:

- Attempt to view the secured item pool prior to testing or to record items as testing occurs. The most common way this threat is realized is by teachers or administrators enrolling pseudo-students so that educators can take the test.

- Retesting students solely to improve a growth score. Since the primary purpose of interim testing is to provide information to help the teacher plan instruction, we believe very strongly that schools should not retest students with valid scores. Threats to validity emerge when retesting procedures are not consistently followed (for example if educators “cherry-pick” retesting to try and inflate their growth scores).
- Inappropriate help or accommodations during the test. Providing help to students that is not specifically authorized in an accommodation.
- Students who are not assigned to teachers within the accountability system.

We recommend that the following safeguards be in place in order to assure the integrity of the testing process, particularly when results are to be used for teacher evaluation. Most of these procedures are not unique to the MAP assessment; they are procedures we would recommend for administration of any test.

- Authorize only one person, preferably a non-teacher, in each school to enroll students (i.e., add student profile data) in the testing system. Most enrollment is done through the central office, but schools need the ability to enter new students into the system, sometimes at the last minute.
- Adopt a policy that prohibits enrolling non-students in the MAP system for the purpose of enabling teachers, parents, or others to take practice tests (thus exposing the item pool) and require that all students have a unique division- or state-issued ID. The division should periodically audit official enrollment records against the record of tested students to confirm that this policy is being followed.
- The MAP system provides data to schools about the validity of individual test events that can be used to ensure the measurement is accurate for each student. These data include test duration, percent correct (on adaptive tests, nearly all students should get approximately half the items right), and when required item-response data. We encourage schools to retest students when these data indicate that an administration is likely to have produced an invalid result. Adopt a policy to govern retesting of students in the MAP system, set specific empirical criteria for testing (for example, all students with test durations under fifteen minutes or percent correct under thirty-five percent should be retested), and require that policy be followed each term. Retesting requests should be approved by the principal, and the division should randomly audit records to assure that retesting policies are being followed.
- Require that accommodations be limited to students with Individual Education Plans and that accommodations allowed be limited to those specifically identified in the plan. Accommodations offered should be noted and approved by the principal for each student.
- Adopt a policy that prohibits teachers from transcribing or saving items from the test and should communicate that policy to teachers.
- MAP testing is normally conducted in a computer lab. We recommend that the test be proctored by both the teacher and a second person (normally the educator or assistant manning the lab). This discourages teachers offering inappropriate help to students. It also has several other benefits. First, the presence of the person responsible for the lab assures that someone in the room is available to troubleshoot problems if they occur. Second, it assures that the teacher has an opportunity to monitor testing conditions to ensure they are appropriate and that the test is properly administered.

- If a computer lab is used for testing, we also encourage recording and preserving test activity on unobtrusive security cameras. This policy helps the division better assure test security, and also provides the testing procedures integrity in the event that a teacher is questioned on this.
- The division should establish procedures to make sure assignments of students to teachers are accurately recorded in order to ensure the integrity of results. These procedures should require that all students are accounted for, and should require validation of teacher assignments by both the teacher and the school principal.

## Section 4: Test Administration Procedures

### Requirement: Section 4: Test Administration Procedures 4.1

#### 4.1

Describe the administration procedures necessary to produce growth scores. For example, is the assessment designed to be administered multiple times during the year or administered once in the fall and once in the spring?

MAP assessments are a measure of both status and growth. Beginning with the 2011 Norms Study, a student's achievement and growth can be referenced from an estimated growth trajectory the student shares with his/her academic peers (i.e., those students who performed equally well when tested during the same instructional week). Given the student's observed RIT score (and its accompanying SEM) from a test occurring at the specific instructional week, the student's normative performance on the test may be directly determined. The new approach may also be used to predict the student's performance if he/she had tested a week earlier, or several weeks later. As a result, performance may be evaluated at alternative and various points on the relevant instructional calendar and, in this sense, is less rigid than the more common test term window designations would suggest. The new procedure facilitates the determination of the Conditional Growth Index (CGI) which enables the user to directly determine the growth by the student's second testing occasion relative to the student's previous performance, and when compared with academic peers.

The depth of MAP for Primary Grades item pools for Virginia will be sufficient for three test administrations per year. NWEA provides four default testing window dates for the purpose of establishing data norms; however, Virginia LEAs will have the flexibility to select test term windows within the following default windows:

- Fall: August 15th through November 30th
- Winter: December 1st through February 28th
- Spring: March 1st through June 15th
- Summer: June 16th through August 14th

While flexible, it is important to establish beginning and ending dates for MAP testing for the purposes of data accuracy and division-level reports. The period of time defined by these dates is known as a test term window. The test term window is useful for setting an end date at which an

**Requirement: Section 4: Test Administration Procedures 4.1**

LEA will declare the testing term complete and subsequently have the option to order division-level summary reports. NWEA reports will only display student results that occurred within a division's defined test term window.

NWEA offers Virginia divisions the following recommendations when selecting test window dates:

- Strive for a three week or less testing window to provide comparisons based on students that have received about the same amount of instruction.
- Keep test window dates consistent from one academic year to the next to allow for valid growth comparison from year to year.
- Establish test window dates using the following recommendations to utilize normative data populated into reports:

Time Category	NWEA Recommendation
Fall test window timing	Beginning of academic year (weeks 1-7)
Winter test window timing	Middle of academic year (weeks 15-21)
Spring test window timing	End of academic year (weeks 28-34)
Time between fall and spring testing	Roughly 32 weeks of instruction
Time between testing in consecutive fall or spring terms	Roughly 36 weeks of instruction

**Requirement: Section 4: Test Administration Procedures 4.2**

4.2

Describe any processes used for pre-identifying and/or registering students for testing. Include what data, including the State Testing Identifier, are collected for each student, how data are collected or transmitted, and how data are maintained and securely managed.

NWEA acknowledges the importance of rostering students prior to testing to ensure the right student receives the right assessment. As a result, NWEA offers comprehensive rostering processes and systems to address partner needs. To ease test administration, NWEA's rostering processes and systems are implemented to support finalizing registration in advance of test sessions. For those cases where a student must be registered immediately prior to testing, both MAP platforms provide a solution to roster individual students. For processing larger roster, or program files, NWEA proposes the following process and systems solution.

## Requirement: Section 4: Test Administration Procedures 4.2

### Web-Based Platform

Using the Administration and Reporting Center interface, division users will have control over the process of importing division-wide or school-wide data files to the web-based platform with student, teacher, class, school, and program information, including the State Testing Identifier. Division users assigned the Data Administrator role can upload complete or partial roster files and programs files to the system. Accommodation needs for individual students can be included in the roster files and will be tracked by the system.

The data files uploaded through the Administration and Reporting Center must adhere to the NWEA roster file and programs file templates; these templates define data fields, character limitations, and other restrictions based on business rules that enable the system to add and update existing information in the system with imported data. The roster files and program files templates are provided in Appendix M. Roster files and program files must be in the CSV (comma-separated values) file format to be imported into the web-based assessment system and can be accepted from any SIS system that allows for configurable data exports.

NWEA's web-based assessment system provides end users the flexibility to import all roster data at one time, or to submit partial data imports. The Assessment and Reporting Center interface accepts the following types of roster imports:

- **Instructor-Only Import** allows instructor access to tutorials, the student warm-up, user guides, and instructional resources, such as DesCartes: A Continuum of Learning, before classes begin. This is especially helpful for new instructor orientation. Divisions might also use this import type to update instructor names and e-mail addresses as needed.
- **Student-Only Import** is convenient for schools that need to prepare students for testing within a very short timeframe when instructor and class data is not yet available. This data can be augmented by updating individual student profiles or importing an updated roster file with more complete data. Note that students imported with minimum information required for testing (student name and grade) will not appear on reports until their profiles include all reporting attributes.
- **Complete Import** prepares student data so that students are ready for testing, organizes data for reporting, and provides instructor access to their students' reports.

Imported files undergo multiple data validation processes. If the system detects errors in the file format (for example, the data file is not in CSV format) or inconsistencies in the data (for example, the same Student ID is assigned to two different students), the system generates an error report. The data administrator corrects the errors in the source system – typically the division's Student Information System (SIS) – and submits an updated file. Updated roster files and programs files can be submitted via the Administration and Reporting Center interface as often as is needed.

### Client-Server Platform

On the client-server platform, the student enrollment process is similar to that of the web-based platform described above, except:

**Requirement: Section 4: Test Administration Procedures 4.2**

- Instructor management is not supported; all user accounts are managed via NTE Admin
- The roster files, also in CSV file format, have slightly different data fields requirements
- The roster files are submitted via the NWEA Reports Site
- File validation is manually performed by the NWEA Technical Support team

**Data Transfer and Storage Security**

NWEA systems' security architecture is based on the underlying security pillars of information security: Confidentiality, Integrity, and Availability. Confidentiality ensures that all NWEA system users can access only data that their permission level entitles them to access. Integrity ensures that unauthorized users are not allowed to tamper or alter any data they do not have access to. Availability ensures that NWEA systems and data are available to authorized users when they need them. The security architecture of NWEA systems is built to provide high security by employing security best principles.

NWEA systems transmit and receive data over the Internet using industry standard 256-bit secure sockets layer (ssl) using a 2048-bit certificate. This approach to secure data transfers is in compliance with the Family Education Rights and Privacy Act ("FERPA") and Children's Online Privacy Protection Act ("COPPA").

**Requirement: Section 4: Test Administration Procedures 4.3**

4.3

Describe all materials needed for test administration and how school divisions will order and obtain sufficient quantities. Include details of test booklets and answer documents for paper/pencil testing (if applicable), test administration manuals, etc. If applicable, identify any test administration materials school divisions would be responsible for supplying locally (manipulatives, copies of test materials, etc).

MAP for Primary Grades assessments are administered on a computer, typically in a school computer lab, therefore no printed test or administration materials are required. See Section 4.4.5 for the MAP for Primary Grades system technical requirements. See Section 4.4.4 for a description of the resources NWEA provides MAP for Primary Grades system users to help understand the administration of the assessments.

**Requirement: Section 4: Test Administration Procedures 4.4**

4.4

Provide examples of the test administration manuals to be used with the assessment(s).

**Requirement: Section 4: Test Administration Procedures 4.4**

NWEA provides multiple resources to help MAP for Primary Grades users understand the general administration protocol that applies to all MAP assessments and test delivery platforms (client-server and web-based). These resources extend to general information about how MAP for Primary Grades tests function and specific “how to” instructions for administering tests on each test delivery platform. Examples include:

- Guidelines for Teacher and Proctor Duties on Test Days, available on NWEA’s Partner Support website at <http://www.nwea.org/support/article/1042/teacher-and-proctor-duties-testing-days> and included with this application as Appendix N.
- MAP for Primary Grades Administration Characteristics of Tests, included in the NWEA Technical and provided as Appendix O.
- MAP for Primary Grades Proctor and Teacher Instructions is available at <http://www.nwea.org/support/article/1187/proctor-and-teacher-instructions-map-primary-grades>.

Additional MAP for Primary Grades administration documentation is available on the NWEA Partner Support website at <http://www.nwea.org/support/article-15>.

**Requirement: Section 4: Test Administration Procedures 4.5**

4.5

Describe all technology requirements related to school personnel managing the administration of tests and to students completing tests if assessments include technology-based delivery. Include the minimum and recommended hardware and software requirements and network requirements for test administration by school personnel and test delivery to students. Include how assessments are hosted (e.g., locally, vendor, 3<sup>rd</sup> party). Provide examples of user interfaces for test administration by school personnel and test delivery to students. Include descriptions or examples of test navigation and any test tools (e.g., calculator, ruler, highlighter) available to students for testing.

NWEA is one of the first providers of computer-based testing in the United States. Its long history in assessment technology puts it at the forefront of test design and test delivery systems. With this experience, NWEA can deliver to Virginia divisions a high-quality and reliable test with either a web-based or client-server platform. NWEA recognizes not all of Virginia’s schools are equipped with the Internet access required to use a web-based product, so NWEA also offers our client-server platform. These platforms have successfully administered over 262,800 tests in the Commonwealth of Virginia during the 2011-2012 school year.

**NWEA’s Web-Based Platform**

The web-based platform is an end-to-end online testing solution with components that comply with the highest standards for data security, transfer speed, and disaster recovery, while providing outstanding up-time during scheduled system availability. The web-based platform is accessible to any computer inside or outside of the division that meets the hardware and software requirements provided in Appendix P. Key components of NWEA’s web-based system are:

#### Requirement: Section 4: Test Administration Procedures 4.5

- **Online Administration and Reporting Center**, providing authorized users role-based access to administration functionality such as: user, student, organization, program, and test data management. Users will also access the Administration and Reporting Center to create and manage student test sessions, access reports, and instructional resources.
- **The Student Testing Center**, a web-based test taking environment where students log in using a standard web browser to take tests that have been scheduled by proctors.
- **Computer-adaptive tests** designed to measure achievement of students as they grow over time. The tests commonly consist of more than forty questions so the system can provide sub-scores that represent the content goals and an overall score for each student.

NWEA plans to add a new capability to our web-based platform to allow for lower grade Internet connections, expected to be available by July 2014. Until the new capability is available, NWEA recommends divisions and schools with lower-grade Internet connections use the client-server platform as an interim solution. NWEA will carefully monitor the progress of high bandwidth internet service as this capacity grows statewide. We will work with the Virginia DOE and the divisions to determine the optimal process to move divisions to our web-based solution.

#### NWEA's Client-Server Platform

The client-server platform is an end-to-end testing solution for administering tests on the TestTaker Client workstation without Internet connection. Four main components comprise the client-server assessment system:

- **Network Test Environment (NTE) folder(s)**. The MAP Assessments are designed to be administered in a local network environment. The NTE is a folder located on the local network that contains agency data, student data, and test package data.
- **NTE Administration Tool (NTE Admin)**. NTE Admin is an installed application used for managing a division's NTE. This program can be used to perform tasks such as adding students to an NTE on a student-by-student basis, moving students from one NTE to another or uploading test results. To use the NTE Admin, you will need an Internet connection and Microsoft .NET Framework 4.0. Since NTE Admin is only used for administrative functions and not for the actual test administration, it is not necessary to have NTE Admin installed on every computer used for test administration. NWEA recommends at least one computer in each school have NTE Admin installed, and that these computers have a high-speed Internet connection.
- **TestTaker Client**. TestTaker is the software application that students use to take NWEA assessments. Proctors will log in to TestTaker to select the test to be administered and the student to whom the test will be administered. Proctors also have access to the Proctor Administration Menu, where they can pause or terminate a test if needed. Running TestTaker requires network access to the NTE server but does not require an Internet connection.
- **NWEA Reports Site**. The NWEA Reports Site is a secure website where division assessment coordinators, school administrators, and teachers can log in to see their students' test results. The Reports Site also includes administrative tools to help the division assessment coordinator manage the testing season, including a secure upload page for Class Roster Files, a page for declaring division testing complete, and a report order page for school and division

**Requirement: Section 4: Test Administration Procedures 4.5**

summary reports.

The client-server user interface for test administration and delivery by school personnel is nearly parallel to that of the web-based platform.

Both the web-based and client-server assessment systems are deployed and operated independently for each Virginia division, requiring no state-level mediation. Technical requirements for both MAP platforms are included as Appendix P.

**Requirement: Section 4: Test Administration Procedures 4.6**

4.6

Describe accommodations available to students with disabilities and limited English proficient students. Include procedures related to the provision of accommodations to eligible students.

The adaptive nature of MAP for Primary Grades assessments make them appropriate for many students in particular subgroups (i.e., students with disabilities, limited English proficient students, etc.). We recognize that student needs and skills within these groups are not homogeneous. In view of this, we focus accommodations on skills and practices used consistently during instruction and assessment. NWEA-approved accommodations pursuant to state and/or division policy are appropriate for an individual student. The following guidance is available when administering MAP to subgroups requiring accommodations:

**General Guidelines for Accommodations**

- Responsibilities: Scribes, page-turners, educational assistants, and other people supporting a student's test must be neutral in responding to the student during test administration. Assistance in test administration must not lead a student to the correct answer. The student's response must accurately represent the student's own choice.
- Reading Tests: Do not read aloud any portion of the reading test to any student.
- Symbols: Do not pronounce or explain math or science symbols.
- Definitions: Do not define any words within test questions.
- Interpreting Results: While many accommodations are recognized, NWEA has not performed studies to confirm how their use affects assessment scores. By definition, accommodations should allow more precise and more valid estimates of true scores. That is, accommodations are intended to "level the playing field" for students with known disabilities. They should not advantage any student relative to any other student at the same true ability level. However, for certain presentation accommodations noted below (2 and 3), we cannot state with a high level of certainty how validity of test scores is affected due to the potential change in the construct being measured. This qualification should be considered when interpreting test scores or using them to make important educational decisions.

**Requirement: Section 4: Test Administration Procedures 4.6**

**Presentation Accommodations**

1. Simplify or clarify directions
2. Native language translation (oral or signing) of test directions  
NOTE: See “Interpreting Results” under the general guidelines above.
3. Native language translation (oral or signing) of test questions (not answer options) for math, science, or language usage tests  
NOTE: See “Interpreting Results” under the general guidelines above.
4. Read or reread aloud the test directions
5. Read or reread aloud the test questions (not answer options) for math, science, or language usage tests only  
NOTE: See “Symbols” under the general guidelines above.
6. Use visual magnification devices or software (for example, MAP is compatible with ZoomText or MAGic)
7. Use auditory amplification devices, noise buffers, or software
8. Use masks to block portion of screen; for example, the student may use a sticky note, index card, or a blank sheet of paper to move down the screen as he or she is reading

**Response Accommodations**

1. Assign scribe to record responses
2. Dictate responses to a scribe
3. Point to responses for a scribe
4. Respond in native language

**Setting Accommodations**

1. Test an individual student in a separate setting
2. Test a small group of students in a separate setting; for example, in a Title I room or counselor’s office
3. Minimize distractions; for example, use a study carrel

**Timing/Schedule Accommodations**

1. Administer test over multiple sessions in a day
2. Administer test over a number of days
3. Administer test at a particular time of day
4. Allow flexible schedule
5. Extend time allotted by proctor or test administrator (All MAP assessments are designed to be untimed.)
6. Allow stop-the-clock breaks
7. Administer at time of day most beneficial to student

**Requirement: Section 4: Test Administration Procedures 4.6**

8. Offer breaks

**Materials or Devices Accommodations**

1. Provide scratch paper
2. Availability of a calculator (when appropriate, it is provided on the screen)

**Miscellaneous Accommodations**

1. Provide drink during testing
2. Provide snack during testing

In addition to these accommodations, MAP for Primary Grades is compatible with third party software to provide capabilities that address the needs of visually impaired students including magnification, highlighting, and color contrast. NWEA will explore MAP for Primary Grades compatibility with third-party software to accommodate the needs of blind students by evaluating third party software compatibility of text to speech and electronic Braille.

The following table identifies which English Language Learner (ELL) accommodations to the MAP assessment are available within the assessment:

- Read directions orally in English
- Allow flexible schedule – test in sessions, over several days
- Extended time
- Provide scratch paper/allow to mark test booklet
- Administer in a small group
- Administer individually

**Requirement: Section 4: Test Administration Procedures 4.7**

4.7

Describe procedures for completed student tests to be submitted for scoring and reporting purposes.

Since the initial release in 2006, MAP for Primary Grades assessments have been computer-scored by the system as tests are administered. Computer scoring allows NWEA to provide preliminary testing results for the student and teacher at the end of each test to identify students needing intervention in a timely manner and to accurately point to instructional learning objects.

Scoring of MAP assessments is based on Item Response Theory (IRT). Student scores are calculated, not only by employing an algorithm that takes into account the students' performance on the test items, but also the difficulty of the items each student is administered. NWEA's system has been rigorously tested to ensure scores rendered from the system are calculated and displayed correctly.

## Section 5 : Scoring and Reporting

### Requirement: Section 5: Scoring and Reporting 5.1

#### 5.1

Describe scoring procedures for all item types and test forms administered, including implemented quality control measures.

MAP assessments are adaptive and employ a common scoring algorithm. During the assessment, a Bayesian scoring algorithm is used to inform item selection. Bayesian scoring for item selection prevents the artificially dramatic fluctuations in student achievement at the beginning of the test that can occur with other scoring algorithms.

Although the Bayesian scoring works well as a procedure for selecting items during test administration, Bayesian scores are not appropriate for the calculation of final student achievement scores. This is because Bayesian scoring uses information other than the student's responses to questions, such as past performance, to calculate the achievement estimate. Since only the student's performance on the day he/she takes a test should be used to give the student's current score, a maximum-likelihood algorithm is used to calculate a student's actual score at the completion of the test.

NWEA uses the Rasch Item Response Theory model to create its vertical scales, called RIT scales. MAP results, reported as RIT scores, relate directly to the curriculum scale in each subject area.

At inception, NWEA adopted Item Response Theory in general and the Rasch model in particular to create its highly stable vertical RIT scales. There is one RIT scale each for mathematics, reading, and language usage. Using the RIT scale to report test results makes it possible to follow a student's educational growth from testing season to testing season and year to year. Please note that each subject area has a unique alignment to the RIT scale; as a result, scores between subjects are not equivalent.

By using Item Response Theory to create the scales and anchoring item difficulty estimates to them, the RIT scales are comparable from one set of items to another, and from one set of examinees to another. This enables comparisons of the scores from different students, or from the same student at different times, even though different sets of test items are administered. This also allows longitudinal comparison of student performance to be made.

#### QA Process

NWEA's dedicated QA staff work together with NWEA's research team to perform a comprehensive review of all reporting data for accuracy and quality. This includes the validation of all business rules and formulas applied when generating results reports for both standard reports provided via the assessment platform and all custom reports or data extracts provided.

**Requirement: Section 5: Scoring and Reporting 5.2**

5.2

Describe the type of reporting provided (e.g, static and/or dynamic, electronic and/or paper-based, item-level, strand-level, and/or test-level scoring). Include approximate timelines for score reports to be available to divisions, how score reports will be accessed and/or obtained, and samples of student, class, school, and division score reports and sample record layouts for electronic data files.

NWEA assessments provide a comprehensive set of computer-based reports, instructional resources, data tools and comparative information sources to gain insight into student growth and proficiency at the individual, classroom, school, grade and division level. System reports and instructional resources are student-centric, research-based, and data-driven. They provide data to inform instruction, evaluate programs, justify budget initiatives, and make key decisions in the classroom, school, and division levels.

These analytical tools include the following reports:

- Student-level reports including the Student Progress Report and the Student Goal-Setting Worksheet provide current and historical student performance. This includes growth and growth projections as well as normative data and Lexile scores to assist students, teachers, and parents select suitable reading texts. Student Progress Reports can be printed by the educator and shared with parents to create a parent/teacher dialog around their student's learning, performance, and growth. To help guide parents through the various components of this report, a special one page annotated version of the report is provided as an accompaniment. The Student Goal Setting Worksheet is especially helpful in allowing parents to see the growth potential of their child.
- Class and teacher reports provide an overview of class performance and detailed information about each student, including: assessment and goal area performance, norm percentile rank, growth, and predictive performance on the summative assessments once a scale alignment study is completed. Teachers can use these reports to differentiate instruction and identify areas of strength and weakness at a group level.
- School-level reports provide information that can help administrators evaluate programs and foster collaboration.
- Grade-level reports provide summary information to help school administrators assess trends, identify areas of strength and weakness, and percentage of students meeting their growth targets
- Division-level reports which help partners assess performance trends by grade and school. These reports present longitudinal data for a particular school and are valuable resources in planning and monitoring school improvement plans.

NWEA uses a user-centric approach to report design. Throughout the product lifecycle, usability testing is incorporated to understand how a teacher or administrator intends to use the reports, and incorporates this mindset in the development of report designs that are further enforced by the experience of NWEA's research teams and education experts. This user-centric approach provides greater benefit to partners through limiting the amount of time required by the user to learn the content of the reports. The information contained within our reports is intended to aid in

### **Requirement: Section 5: Scoring and Reporting 5.2**

swift and decisive action plans. The MAP assessment system scores each test as it is administered and provides preliminary results for the student and teacher at the end of each test. Reports are generated and accessed online.

Longitudinal Results: NWEA collects and stores student MAP assessment results across time, providing educators with longitudinal data to compare student growth over time. Because MAP results are reported on vertical scales called the RIT scales, it is possible to compare individual students' results over time and compare sets of results. Furthermore, educators can access reports for the current academic year and the previous academic year from NWEA's Reports Site (see page 5 of Appendix Q for a sample ).

NWEA's reports were developed with a student's education team in mind. To help educators make the best use of student test results, NWEA offers a robust menu of easily accessible web-based reports that include student, class, school, and division test results.

MAP for Primary Grades reports give teachers insight into the capacity of their students in areas of mathematics and reading. These reports provide teachers with student performance data that can help to guide strategy lesson plans. MAP for Primary Grades Skills Checklist and screening tests measure students' mastery of foundational skills and grasp of concepts, allowing teachers to make the most of classroom time.

The MAP for Primary Grades Teacher reports not only enable teachers to view overall class results, but assist educators with instructional planning, monitoring student achievement, diagnosing student strengths and weaknesses, and planning individual instruction. MAP for Primary Grades Teacher report results are displayed as colored bands signifying levels of high, medium, and low performance. Class performance results include both unique skills and concepts measured by specific tests. To view a selection of NWEA's sample reports and guidance on how to interpret these reports, see Appendix DD, Annotated MAP for Primary Grades Reports.

Primary Grades Instructional Data: MAP for Primary Grades tests present concepts and skills in emerging literacy and numeracy organized by goal and sub-goal categories in Primary Grades Instructional Data statements. Teachers can use this tool to translate a wide range of student scores into instructional objectives that allow them to focus on individual student learning needs.

#### **Availability and Format of Reports**

The MAP assessment system scores each test as it is administered and provides preliminary results for the student and teacher at the end of each test. Reports showing test results are available to teachers and administrations via login at the web-based reporting interface.

For the Web-based MAP system, reports are accessed via the MAP Administration and Reporting Center. Please refer to the table below for complete information regarding reports availability. Reports can either be viewed online and/or using a PDF viewer, such as Adobe Reader<sup>®</sup>. From the PDF viewer reports can be viewed, saved, or printed.

**Requirement: Section 5: Scoring and Reporting 5.2**

Available Immediately	Available within 24 hours*
<ul style="list-style-type: none"> <li>▪ Class Breakdown by RIT Report</li> <li>▪ Class Breakdown by Goal Report</li> <li>▪ Class by Projected Proficiency</li> <li>▪ MAP for Primary Grades Sub-Skill Performance Report</li> <li>▪ MAP for Primary Grades Skills Checklist Student Report</li> <li>▪ MAP for Primary Grades Skills Checklist Class Report</li> <li>▪ Primary Grades Instructional Data</li> <li>▪ Student Goal Setting Worksheet</li> </ul>	<ul style="list-style-type: none"> <li>▪ District (Division) Summary Report</li> <li>▪ Student Progress Report</li> <li>▪ Class Report</li> <li>▪ Grade Report</li> <li>▪ Achievement Status and Growth Reports</li> <li>▪ Student Growth Summary Report</li> <li>▪ Projected Proficiency Summary Report</li> </ul>

In the client-server MAP system, reports are accessed via the NWEA Reports Site. Please refer to the table below for information about reports availability timeframe and format.

Available Reports	Timeframe	Format
Teacher- and class-level reports	The day after test results are uploaded to NWEA	PDF*
Grade-, school-, and district-level (division-level) reports	Within 24-72 hours after requesting reports	Crystal Reports*, tab-delimited (Data Text File), and CSV (Comprehensive Data Files)
Dynamic Reports showing class-, school-, grade-, and district-level (division-level) growth and proficiency data	One week after test results are uploaded to NWEA	Online display; can be exported to PDF

**Report Samples**

Samples of student, class, school, and division reports are included in Appendix DD, Annotated Reports. In addition to sample reports, Appendix DD, provides guidance on how to interpret these reports. The annotated reports display sample versions along with helpful explanations of what each of the different report components mean. NWEA partners use the Annotated Reports to gain better understanding for how to interpret assessment scores. Additionally, NWEA offers onsite workshops and online training specifically designed to help educators interpret results and adjust instructional practices to meet their students’ academic needs. See Proposal Narrative, Section a. for information about NWEA’s online support and professional development workshops.

**Requirement: Section 5: Scoring and Reporting 5.3**

5.3

Describe all data tools available to school division staff for the analysis of data and the creation of customized reports.

NWEA is committed to providing accurate and timely results to Virginia educators. In addition to offering the aforementioned menu of reports, NWEA also makes available two additional options for obtaining data: raw data files and custom reporting.

**Raw Data Files**

For divisions, the MAP system generates a Comprehensive Data File that contains assessment results, student data, class assignments, and special program (sub-group) designations in a flat text format (common-separated values (CSV)) for further analysis and use in external data or reporting systems, i.e. data warehouses. See Appendix M for file format sample.

**Optional Division Custom Reporting**

If Virginia divisions require custom reporting capabilities, NWEA is able to deliver sophisticated custom reports designed alongside our partners at a negotiated rate.

- 
- <sup>1</sup> Wang, S., McCall, M., Jiao, H., & Harris, G. (April, 2012). Construct Validity and Measurement Invariance of Computerized Adaptive Testing: Application to Measures of Academic Progress (MAP) Using Confirmatory Factor Analysis. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.
- <sup>2</sup> Wang, S., Jiao, H., & Zhang, L. (2012, April). Validation of Longitudinal Achievement Constructs of Vertically Scaled Computerized Adaptive Tests: A Multiple Indicator Latent Growth Modeling Approach. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada.
- <sup>3</sup> Kingsbury, G. G., & Wise, S. L. (in press). Turning the page: How smarter testing, vertical scales, and understanding of student engagement may improve our tests. In R. W. Lissitz and H. Jiao (Eds.) *Computers and their impact on state assessment: Recent history and predictions for the future*. Charlotte, NC: Information Age Publishing, Inc.
- <sup>4</sup> Wise, S. L. (2012). *The utility of adaptive testing in addressing the problem of unmotivated examinees*. Manuscript submitted for publication.
- <sup>5</sup> Wise, S. L., Kingsbury, G. G., & Hauser, C. (2012). *How do I know that this score is valid? The case for assessing individual score validity*. Manuscript submitted for publication.
- <sup>6</sup> Wise, S. L., Ma, L., & Theaker, R. A. (2012, May). *Identifying non-effortful student behavior on adaptive tests: Implications for test fraud detection*. Paper presented at the Conference on the Statistical Detection of Potential Test Fraud, Lawrence, Kansas.



## Copies of Proposed Assessment

Because NWEA's assessments are computer-adaptive, paper-based copies of MAP for Mathematics, Reading, and Language Usage, MAP for Science, and MAP for Primary Grades assessments cannot be provided as part of this proposal. However, to enable the evaluation panel to examine to our assessments, access to our web-based assessments is provided below.

Note: this is a sample site populated with generic student data. Additionally, the assessments accessible via the sample site are not aligned to the Virginia Standards of Learning.

### Instructions for Accessing Sample Administrative and Student Site

NWEA has set up sample "sandbox" administrative and student sites for the Virginia evaluation panel to use the MAP assessment program. Instructions on how to access the site and set up a test event as a Proctor are provided below.

This sample site is a shared environment, meaning multiple users have access to the information added to the site. For this reason we recommend that real user information not be added. Sample data has been added for the Virginia evaluation panel which is a combination of generic and fictitious names.

**Generating Reports:** A small number of test events were generated for users to view real-time reports. Please note the data volume is low and that a more robust view of reports may be found in the Reports Reference documents. When generating reports, use:

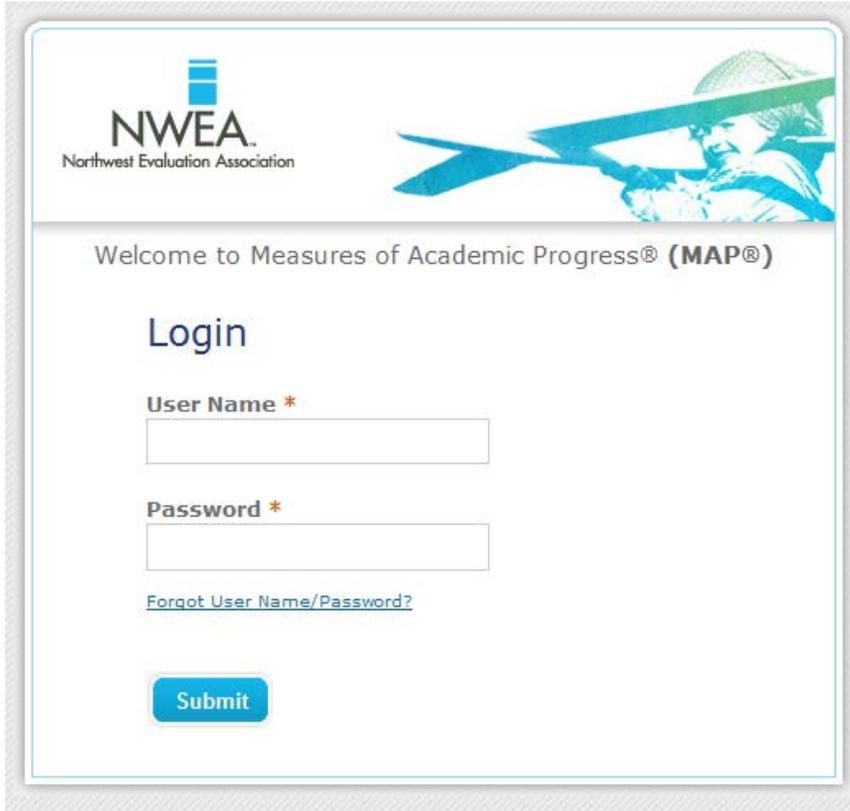
- Term: Winter 2011-2012
- School: Sample School 1
- Instructor: Teacher10, Teacher10

**Help and Tutorials:** The tutorials on the homepage serve as a good resource for learning about the system. If you have questions at any point, please use the "Help" link in the upper right corner of the page.

#### *Sample Administrative Site*

To access the sample administrative site, point your browser to the following address:

<https://samplepartner6-admin.mapnwea.org/admin>.



**Figure 1: The MAP Login Page.**

This link will take you to the MAP Login page. At the prompt enter the following:

- User Name: Virginia
- Password: virginia1

**When you have successfully logged into the MAP sample site, you will land on the following page. From here you may:**

- See what it takes to schedule and administer a test from a teacher’s perspective
- View achievement reports at the teacher, school, and division level
- View growth reports at the teacher school, and division level
- View linked instructional materials or content

#### *Sample Test Event*

To take a student test, first initiate a test event on the site. The following directions will provide users with the experience of a Proctor. To start a test, follow these instructions:

1. From the left-hand column on the MAP home page, select “Manage Test Sessions.”

Copies of Proposed Assessment

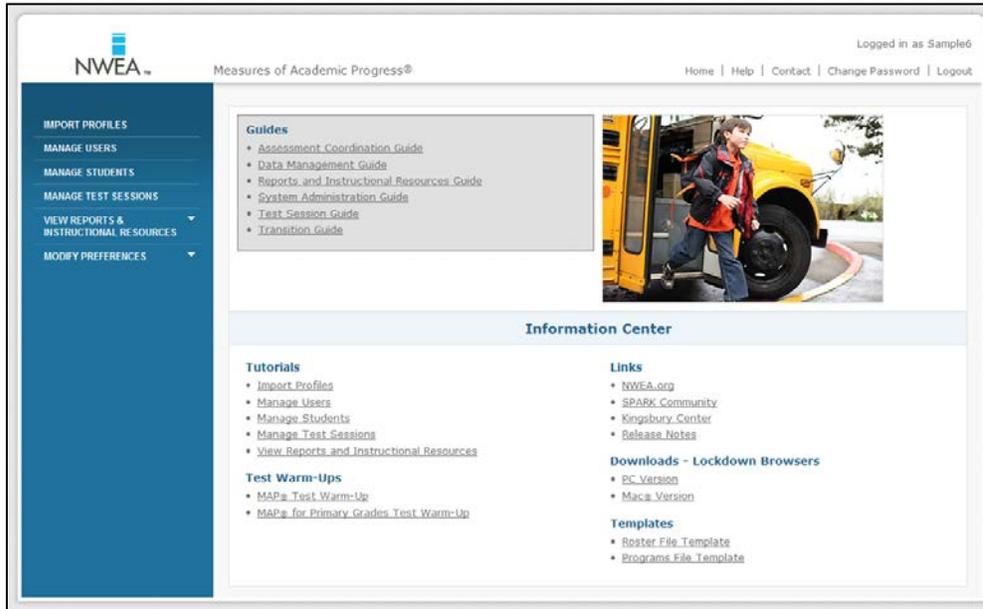


Figure 2: The MAP Home Page.

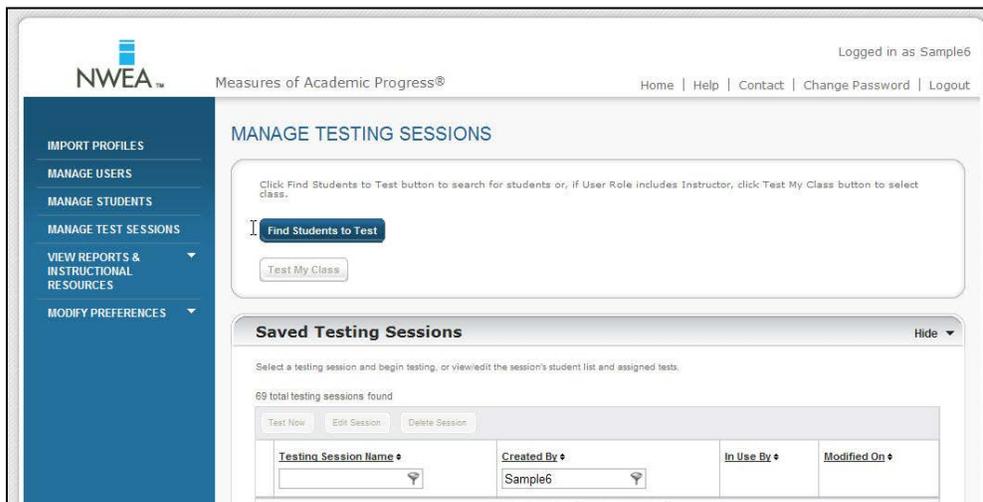


Figure 3: Manage Testing Sessions Page.

2. In Manage Testing Sessions, click the blue button to “Find Students to Test.”

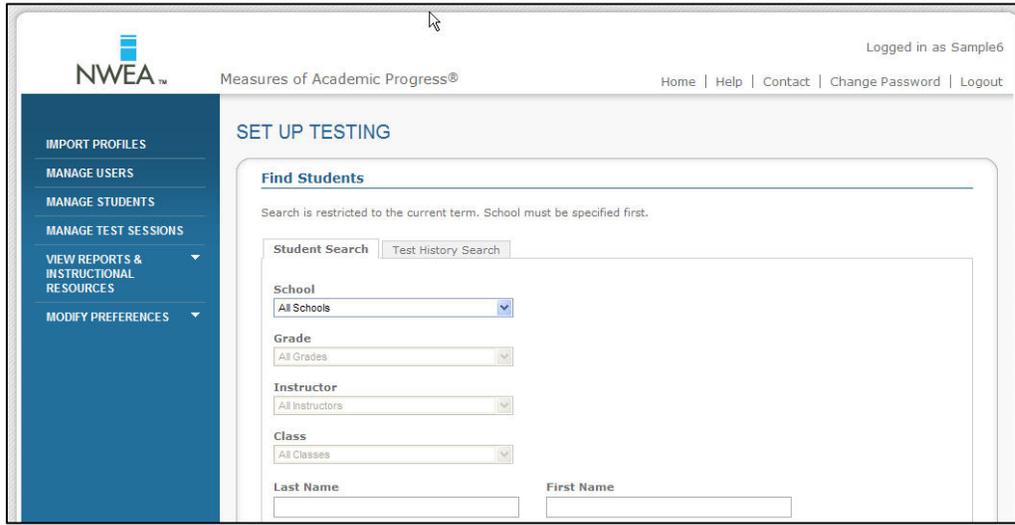


Figure 4: Find Students to Test Page.

3. From the “Set Up Testing” page, use the drop-down menus to select a school, grade, instructor and class. This selection doesn’t have to be specific. Click “Search” and a list of student names will appear. Click “Add students.”

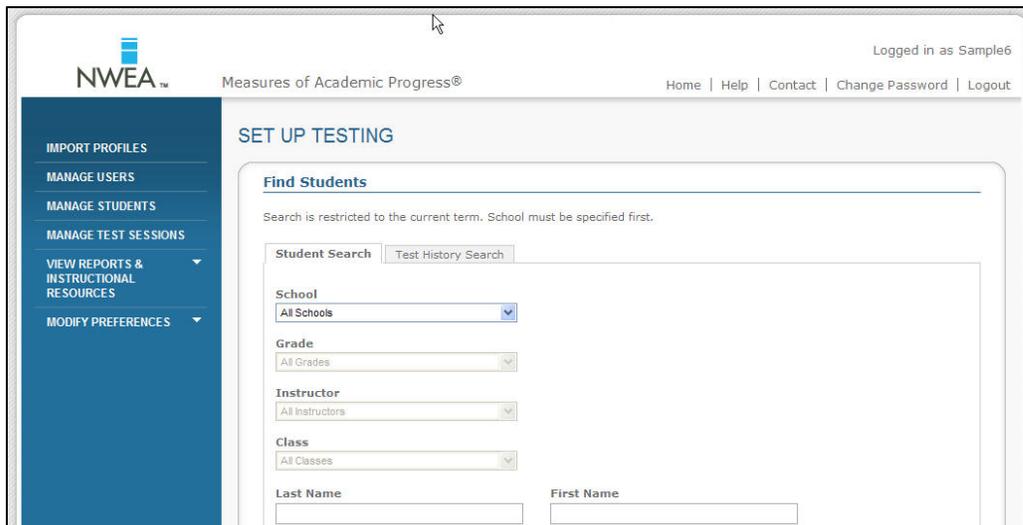


Figure 5: Set Up Testing Page.

4. The names of the students will be added to the bottom of the Set Up Testing page. Then, click “Start Test.”
5. The system will provide a Test Session Name and Password. Write these down.
6. Open a new browser window and go to: <https://sampleprd-test.mapnwea.org/test>.
7. Use the Test Session Name and Password into the login page and hit “Join.” This will launch the student test.

**Join Session**

Welcome to Measures of Academic Progress (MAP).

Please enter your test session name and password to join your test session.

**Test Session Name**

**Password**

**Join**

**Figure 6: Join Session Login Page.**

If you are having issues with the sample site and wish to contact an NWEA technical consultant, please email [techconsultantgroup@nwea.org](mailto:techconsultantgroup@nwea.org) for immediate support.



(This page left intentionally blank)

## Continuity of Services

a) The Contractor recognizes that the services under this contract are vital to the Local Education Agencies (LEAs) and must be continued without interruption and that, upon contract expiration, a successor, either the LEAs, the Agency, or another contractor, may continue them. The Contractor agrees:

(i) To exercise its best efforts and cooperation to effect an orderly and efficient transition to a successor;

(ii) To make all LEA data available to any successor at an appropriate time prior to the expiration of the contract to facilitate transition to successor; and

(iii) That the Agency Contracting Officer shall have final authority to resolve disputes related to the transition of the contract from the Contractor to its successor.

b) The Contractor shall, upon written notice from the Contract Officer, furnish phase-in/phase-out services for up to ninety (90) days after this contract expires and shall negotiate in good faith a plan with the successor to execute the phase-in/phase-out services. This plan shall be subject to the Contract Officer's approval.

c) The Contractor shall be reimbursed for all reasonable, pre-approved phase-in/phase-out costs (i.e., costs incurred within the agreed period after contract expiration that result from phase-in, phaseout operations) and a fee (profit) not to exceed a pro rata portion of the fee (profit) under this contract. All phase-in/phase-out work fees must be approved by the Contract Officer in writing prior to commencement of said work.

NWEA will provide efficient and orderly transition services for any LEA that chooses to discontinue using NWEA's assessments or at the end of the contract period (including contract extensions), in accordance with NWEA Program Management procedures. NWEA will transition all material related to an LEA's implementation of the MAP assessments that does not violate NWEA's regulations on intellectual property, trademarks, or copyrights to the succeeding vendor.

NWEA will provide a phase-in and phase-out services for up to ninety days after of receiving written notice from the Contract Officer that the contract has expired. The phase-in and phase-out services proposed by NWEA will be approved by the Contract Officer prior to implementation of the plan. NWEA agrees to be reimbursed for all reasonable, pre-approved phase-in/phase-out costs approved by the Contract Officer.



(This page left intentionally blank)