**STUDENT ASSESSMENTS AND ASSOCIATED GROWTH MODELS FOR TEACHER AND PRINCIPAL EVALUATION**

**FORM *C***

# PUBLICLY AVAILABLE SERVICES SUMMARY

This form will be posted on the New York State Education Department's Web site and distributed through other means for all applications that are approved in conjunction with this RFQ to allow districts and BOCES to understand proposed offerings in advance of directly contacting Assessment Providers regarding potential further procurements.

| **Assessment Provider Information** | |
|---|---|
| Name of Assessment Provider: | The Northwest Evaluation Association (NWEA) |
| Assessment Provider Contact Information: | Michelle LaPlatney, NWEA Account Executive<br>Ph: (973) 769-9985<br>Email: michelle.laplatney@nwea.org<br>121 NW Everett Street<br>Portland, OR 97209 |
| Name of Assessment: | Measures of Academic Progress (MAP) |
| Nature of Assessment: | ☐ ASSESSMENT FOR USE WITH STUDENT LEARNING OBJECTIVES WITH A TARGET SETTING MODEL; OR<br><br>☒ SUPPLEMENTAL ASSESSMENT WITH AN ASSOCIATED GROWTH MODEL:<br>☐ GAIN SCORE MODEL<br>☐ GROWTH-TO-PROFICIENCY MODEL<br>☐ STUDENT GROWTH PERCENTILES<br>☐ PROJECTION MODELS<br>☒ VALUE-ADDED MODELS<br>☐ OTHER: |
| What are the grade(s) for which the assessment can be used to generate a 0-20 APPR score? | Grades K – 2 |
| What are the subject area(s) for which the assessment can be used to generate a 0-20 APPR score? | Mathematics and Reading |
| What are the technology requirements associated with the assessment? | All NWEA reports for MAP assessments are available online through the MAP Administration and Reporting Center (MARC), which provides a comprehensive set of intuitive web-based reports, instructional content, data |

| | |
|---|---|
| | tools, and comparative information sources. The center is accessible to educators and administrators from any location with an Internet connection, at any time outside of scheduled maintenance. Maintenance typically occurs once per month on the weekends, and NWEA provides an advance schedule to partners.<br><br>For current technical requirements, please see: https://teach.mapnwea.org/impl/QRM2_System_Requirements_QuickRef.pdf<br><br>MAP is also supported for current partners who have been using our client server platform. Technical requirements for this platform are unchanged from our previous application. However, all new partners will implement MAP assessments via the web-based platform described above. |
| Is the assessment available, either for free or through purchase, to other districts or BOCES in New York State? | ☒ YES<br><br>☐ NO |

**Please provide an overview of the assessment for districts and BOCES. Please include:**

- **A description of the assessment;**
- **A description of how the assessment is administered;**
- **A description of how scores are reported (include links to sample reports as appropriate);**
- **A description of how the Assessment Provider supports implementation of the assessment, including any technical assistance. (3 pages max)**

**Overview of MAP Assessments for Grades K – 2**

The Northwest Evaluation Association™ (NWEA™) is pleased to provide this introduction to the Measures of Academic Progress® (MAP®) assessment system for grades K – 2. Our web-based, item-level adaptive interim assessments measure student achievement in mathematics and reading in grades K – 2 (MAP for Primary Grades). We offer our experience, expertise, and research-based assessments to districts and BOCES in New York to empower educators to accurately measure student achievement and growth within the school year and across years. The precise data MAP assessments provide, aligned to New York State P – 12 Common Core Learning Standards (NYSCCLS) will allow educators to make the kinds of immediate instructional decisions that can effect positive change for every student.

Unlike traditional, paper-and-pencil standardized assessments, MAP assessments for grades K – 2 are able to tailor item selection to the ability levels of examinees. The adaptive nature of our grades K – 2 assessments ensures that students are assessed on material appropriate to their individual level. Further, the test items are designed with young learners in mind. It is engaging for early learners to click and move pictures around the screen, similar to what they do in computer-based educational games. MAP assessments for grades K – 2 also provides test warm-ups that allow students to quickly become familiar with question types before the test even starts.

MAP assessments for grades K – 2 meet the unique needs of early learners by providing appropriate supports for students in the primary grades. These supports include displaying interactive elements and providing interactive items with audio support, to ensure beginning readers understand the tasks presented by the assessment. Many items on the MAP tests are interactive in nature, meaning students can manipulate and construct answers based on the learning being assessed.

MAP assessments for grades K – 2 accurately reflect each student’s instructional level and provide educators immediate, highly accurate, detailed data about what students know and what they are ready to learn next. MAP assessments are recognized as one of the highest quality, research-based interim assessment systems available due to the strength of the vertical scale, our test reliability and validity, adaptive test algorithms, large item pools, and the stringent item development processes we follow. Our system of support includes robust professional development services, account management services, and technical support to ensure educators can administer the tests easily, understand the results, and take action to improve instruction.

As one of the founders of the adaptive testing movement, NWEA has over thirty years of experience in this field, and well understands the value of providing rich and efficient testing experiences that are tailored to the individual learning of each student.

MAP assessments are recognized as one of the most accurate measures of student achievement and growth in the market, and are used by over 7,400 NWEA partners including state departments of education, school districts, private schools, charter schools, foundations, universities, school reform groups, the Bureau of Indian Education (BIE), and international schools.

With the MAP assessment system, educators also gain:

- A stable, grade-independent, vertical scale that measures growth with precision even as standards and education continue to change
- Powerful reporting options for multiple stakeholders, including district leadership, principals, teachers, students, and parents
- Fast access to data, as MAP produces student scores immediately after test completion
- Longitudinal data to track student growth over time
- Growth and achievement norms that allow for the national comparison of the achievement and growth patterns of students in your district or BOCES with students in all fifty states
- Experienced implementation, technical support, and account management personnel to ensure smooth onboarding and administration of assessments

- Secure and reliable data available for export to student information systems and data warehouses
- A stable testing platform that can reliably deliver assessments and results, even in large districts; the platform supports 120,000 concurrent test events with item response times within milliseconds

***MAP Assessment Administration***

We administer more than forty million MAP assessments annually throughout the world. Our assessments have been used to target and improve classroom instruction and advance student achievement across the state, across the nation, and across the globe.

MAP assessments are adaptive, meaning that each test is designed to optimally engage each student by adjusting to his or her instructional level, at the item level, through our adaptive test engine. The assessment begins by delivering the student a grade-level question. If the student answers the question correctly, the test taker is rewarded with a more difficult question. Conversely, an incorrect response triggers the delivery of an easier second question.

All students take a unique version of the test, calibrated to a difficulty level where they will achieve approximately fifty percent correct answers. As a result, struggling students who typically become frustrated during testing, and high achievers who may find traditional tests boring, encounter a test that is appropriately challenging, which increases engagement and reduces the propensity to guess answers at random.

By creating a unique test for each student, educators receive highly accurate information about their students. The adaptive nature of MAP assessments for grades K – 2 helps to capture a true portrait of the individual student's achievement level. Rather than simply indicating what a student might be able to do relative to grade-level standards, MAP tests indicate what a student is ready to learn relative to the NYSCCLS – not bound by grade.

Our partners' experience of the MAP assessment system is that it is intuitive, easy to operate, and engaging for both students and educators. The MAP system has a visually appealing interface that is simple to navigate, providing users with embedded, page-specific online help, guides, and tutorials on-demand. All of this supports our ability to keep educators' time invested in learning the application to a minimum, while maximizing the ability to obtain useful and actionable information from the data.

***Fast, User-Friendly Reporting***

Upon completion of a MAP test, the assessment software calculates each student's score and immediately displays the score for the subject and goal areas via the end-of-test screen. While student reports are available immediately after testing, reports that aggregate data at the classroom, school, and district level become available after testing for that group of students is complete. All reports are available online using the MAP Administration and Reporting Center (MARC) and can be accessed from any location with an Internet connection. A review of the reports available in the MARC can be found here: https://www.nwea.org/assessments/map/reporting-data/featured-reports-measures-academic-progress-map-map-primary-grades/.

For more information on the reporting offered by the MAP assessment suite, please visit our reporting resources page with Quick Facts, a short video overview of standard MAP assessment reporting, and additional information on customized reports at: https://www.nwea.org/assessments/map/reporting-data/.

***Support Services***

MAP assessments are a fully hosted solution and are easy to implement and to administer. We also provide outstanding support to partners throughout our relationships to make sure we are effectively meeting their needs. The NWEA Partner Accounts and Partner Services teams are responsible for providing our partners with account management, professional development, implementation, and technical support. A culture of continuous improvement coupled with hiring practices focused on selecting individuals with strong customer-service orientation has created an excellent team that is well respected by our partners. Our staff is well-versed in the benefits and challenges inherent in the implementation and ongoing delivery of computer-based adaptive assessment systems.

**Implementation Services**

During a partner's first testing season, NWEA assigns an Implementation Support Specialist to proactively guide designated 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.

After a partner has completed their initial implementation of the MAP system, ongoing account management duties are transitioned from the Implementation Support Specialist to the Account Manager.

**Account Management**

The designated Account Manager is the point of contact at NWEA for any partner questions. Our Account Managers work out of our national headquarters in Portland, Oregon, or out of their remote home offices located around the country, and are available via phone or email on a daily basis. Most inquiries receive a response within twenty-four to forty-eight hours. Account Managers schedule periodic check-in meetings with partners to answer questions, follow up on any open issues, schedule additional professional development as needed, and collaborate with staff on their plans for the next season of testing.

**Technical Support**

Our highly skilled Technical Support Team is available by toll-free Support line, email, or chat Monday through Friday, 7:00 a.m. to 8:00 p.m. Eastern Time (EST), excluding federally recognized holidays observed by NWEA. Our Technical Support staff is well-versed in the implementation of our assessment systems and can provide assistance with generating roster files, configuring system components, accessing online reports, and answering any questions that arise in the use of the systems.

**Reciprocal Partner Communications**

Incoming calls, emails, and chats are routed through a tiered support system for effective triaging based on the nature and urgency of the question or issue. Our call routing and escalation processes efficiently route service requests to the appropriate personnel. NWEA staff respond with accurate, timely, courteous, and consistent service. To optimize responsiveness and maintain consistently high customer satisfaction, we use established Information Technology Infrastructure Library-based (ITIL) escalation protocols.

Technical Support staff log all partner inquiries and the resolution in a Customer Relationship Management (CRM) application, providing an historic record of all partner support interactions. The management team regularly reviews these log reports to identify trends, escalate bug fixes, and analyze options for enhancements to be included in future product releases.

NWEA also provides product release notes for our partners to communicate new product features and other "under the hood" changes that improve system performance and stability. These monthly Partner Update newsletters are hyperlinked from the MARC, and also provide a synopsis of upcoming partner-facing changes that explain what changes or updates will be taking place in the coming days. Prior to the release, partners receive a partner update email that highlights important partner-facing changes.

**Please provide an overview of the student-level growth model or target setting model for SLOs for districts and BOCES, along with how student-level growth scores are aggregated to the create teacher-level scores, and how those teacher-level scores are converted to New York State's 0-20 metric.**

NWEA is partnering with researchers at Education Analytics Inc. (EdAnalytics) to provide districts in New York with value-added estimates of educator effectiveness. The use of this approach is the best means to ensure that student results accurately capture the contributions educators made to the growth of their students, by eliminating noise from factors external to the educator and outside his or her control that could lead to mistaken generalizations about that individual's performance.

EdAnalytics has an established value-added process that yields estimates of the contribution of educational units (classrooms and schools) and agents (teachers and principals) to student achievement or other student outcomes, after controlling for external (non-school) sources of student achievement growth, including prior student achievement and student and family characteristics. EdAnalytics' objective is to facilitate valid and fair comparisons of productivity with respect to student outcomes, given that schools often serve very different student populations.

EdAnalytics collects data files from a variety of sources including the NWEA test files, districts' SIRS data files, and templates especially created for this work. By combining all these files, EdAnalytics can gain a comprehensive view of both school and non-school factors that may have affected the growth of a particular teacher's students. EdAnalytics' methodology works by comparing the growth of each student to observationally similar students (students with similar starting points and characteristics). If students in a teacher's classroom tend to grow faster than observationally similar students, the teacher receives a high value-added rating, meaning the value that teacher added to student growth was greater than expected.

EdAnalytics produces results for each teacher that indicates how a teacher's students grew relative to the average of similar students. These results are then translated to the proposed 0-20 HEDI scale. In EdAnalytics' proposed system, a teacher whose students grew at a typical rate for similar students will receive a "sixteen", the middle of the "Effective" range. Results below sixteen do not indicate that students lost knowledge, but rather indicate that student growth was below-average as compared to the growth of similar students.

In order to create and update its methodology throughout the course of this existing work, EdAnalytics has employed its "co-build" approach where stakeholders from participating districts participate in New York MAP Value-Added Advisory Council meetings. At these meetings, stakeholders from participating districts along with experts from EdAnalytics and NWEA evaluate the ongoing development of the growth metrics and ensure the methodology used continues to address evolving policy needs while remaining compliant with state regulations.

| **New York State Next Generation Assessment Priorities**<br>Please provide detail on how the proposed supplemental assessment l or assessment to be used with SLOs addresses each of the Next Generation Assessment Priorities below. | |
|---|---|
| **Characteristics of Good ELA and Math Assessments (only applicable to ELA and math assessments):** | The MAP assessments are consistent with many of the criteria supplied by the Achieve the Core Assessment Evaluation Tool (AET). However, the adaptive and across-grade nature of the MAP assessments mean that some of AET criteria do not apply since those criteria focus on within grade assessments. The MAP assessments are specifically designed to cross grades as this structure allows the assessments to measure where each individual student is performing, show growth, and provide teachers with more precise information about what their students know.<br><br>The AET criteria for ELA tests in grades 3 – 12 include: the importance of the complexity and quality of the texts, test questions that are standards-based, texts that reflect the distribution of text types and genres required by the state standards, vocabulary items that assess words in context and focus on central ideas in the text, items around conventions and writing strategies that focus on the standards and actual practice, a variety of items types that are appropriate to the standards, and test blueprints that reflect the standards. The AET for ELA tests does not specifically mention K – 2, however MAP assessments in the K – 10 grade range adhere to the information provided below.<br><br>The items used in the MAP assessment item pools are all aligned to the NYSCCLS and the test blueprints (goal structures) are built to reflect the organization of the NYSCCLS and provide evidence of alignment to the standards. For example, MAP for Reading assessments available for grades K – 10 cover Vocabulary, Literary Texts, and Informational Texts explicitly. |

| | |
|---|---|
| | Recognizing that some Standards do not lend themselves to selected-response items, NWEA continues to add alternate item types such as:<br>▪ Technology-enhanced items: interactive items that students can manipulate to construct answers based on the learning assessed<br>▪ Drag and Drop: a student drags one or more objects (e.g., numbers, words, pictures) from one location of the screen to another to create an answer<br>▪ Click and Pop: a student clicks on one or more answer objects (e.g., numbers, words, pictures) that automatically move to a pre-selected location on the screen<br>▪ Hot Spot: a student clicks on one or more objects (e.g., geometric shapes, text, symbols) to select answer option(s); items indicated by the student are highlighted<br>▪ Common stimulus reading items: sets of items associated with a single literary or nonfiction stimulus (e.g., an extended passage). Students read the passage and answer a series of selected-response items that target a variety of skills, requiring students to engage different cognitive processes. Extended passages allow for a more authentic and sustained reading experience, where students can respond holistically to a complete text of appropriate rigor<br>▪ Items containing video and animation<br>▪ Items that use drawing and orienting lines and figures<br>▪ Virtual performance tasks<br>▪ Constructed-response items<br><br>All NWEA items go through a rigorous item development and review process. The process yields items with strong alignments to the breadth and depth of the NYSCCLS. In order to achieve this, we have developed a deep understanding of the standards and use a variety of approaches and item types to |

| | |
|---|---|
| | assess them. Items are developed and reviewed through a variety of lenses, including how they align to the targeted standard and grade level, how they adhere to the principles of Universal Design, and whether they are free from potential bias and sensitivity issues. Additionally, the literary and informational texts used by items are evaluated both quantitatively and qualitatively, looking at criteria such as knowledge demand, language conventionality, and clarity in order to determine the readability and complexity of each passage.<br><br>The AET criteria for Math tests in grade K – 12 include some criteria that assume an on-grade-level, fixed-form assessment: a focus on the concept of Major Work for the grade, items should not assess topics before they are introduced in the standards, reflect the grade-by-grade progression of the standards, and score reporting that reflects the emphasis of the grade. NWEA assessments are adaptive tests that are designed to assess students where they are regardless of grade level and to show growth regardless of grade level. Students will see items that are aligned to standards above or below their grade level. However, NWEA's RIT scores and Learning Continuum reports make it simple to determine where students are performing with regard to grade level standards. They also make it possible to see what skills and concepts students who are performing above or below grade level know. The MAP assessments for mathematics do contain items that align to the standards and therefore reflect the three aspects of rigor and the connection between content standards and practice standards.<br><br>In addition to specifics about the ELA and Mathematics assessments, AET provides guidance about indicators of a quality assessment. As mentioned above, all items go through a multi-stage item development and review process to ensure high quality items. NWEA assessments are built to be student worthy assessments. The purpose of the tests, the data and reporting needs, and |

| | |
|---|---|
| | learning targets are all considered as part of the test development process. Items are then aligned to the learning targets and go through a field testing and calibration process to place them on the measurement scale and to ensure accuracy and validity. Test designs are carefully constructed to include the relevant content, technical, and psychometric information needed for test construction. The tests and items that result from these defined processes offer valid content, reliability in terms of valid data for students at all levels and across years, and fairness across student populations.<br><br>The items presented to a student in any given test event are determined by the individual student's achievement level and by the test's goal structure. These goal structures group all assessable standards into goal areas that represent content domains and sub-goals that represent common groupings of grade level expectations, which cover related topics along the learning continuum within each standard. Each student is administered a balanced number of items in each goal area to provide an overall score for the content area (reading or mathematics), as well as goal area scores.<br><br>Because MAP tests are adaptive and designed to provide data about students across the achievement continuum – including students who are performing below level or above level – the item pools that support these tests are very large and include items that may range in complexity from the most basic "building block" aspect of a skill to analytical or evaluative aspects of the skill. MAP assessments are designed to assess students where they are, regardless of grade level. However, the MAP assessment for mathematics does contain items that align to the standards and therefore reflect the three aspects of rigor and the connection between content standards and practice standards. |

| | |
|---|---|
| | In addition to specifics about the ELA and Mathematics assessments, AET provides guidance about indicators of a quality assessment. As mentioned above, all items go through a multi-stage item development and review process to ensure high quality items. NWEA Assessments are built to be student worthy assessments. The purpose of the tests, the data and reporting needs, and learning targets are all considered as part of the test development process. Items are then aligned to the learning targets, and go through a field testing and calibration process to place them on the measurement scale and to ensure accuracy and validity. Test designs are carefully constructed to include the relevant content, technical, and psychometric information needed for test construction. The tests and items that result from these defined processes offer valid content, reliability in terms of valid data for students at all levels and across years, and fairness across student populations. |
| **Assessments Woven Tightly Into the Curriculum:** | NWEA believes that each and every student matters, and we offer assessments designed to help guide meaningful classroom instruction. MAP assessments offer a personalized experience for students by adapting to each student's learning level – precisely measuring student progress and growth for each individual. Assessments are designed to be completed within a short amount of time (forty to sixty minutes per domain) and to provide teachers with robust information within twenty-four hours about what each student knows and is ready to learn, which can be used to inform classroom instruction.<br><br>MAP assessments provide teachers with a means to measure the growth and progress of every student over time, regardless of whether a student is performing on, above, or below grade level. In addition, the assessments compare and predict student achievement and growth over time via NWEA achievement and growth norms. These data can be used by teachers to personalize instruction quickly for 1:1, small group, or |

| | |
|---|---|
| | whole class activities. Teachers can also use the data to support efforts to engage students in achieving personalized learning goals and progress via student and family goal setting activities.<br><br>MAP assessments include our proprietary interactive tool for teachers, the Learning Continuum. Teachers can use the Learning Continuum's information to streamline instructional planning, differentiate instruction for both individual students and skill-based activity groups, and better engage students in their learning. It is a powerful shortcut to understanding which skills students are ready to learn.<br><br>Within the Learning Continuum, learning statements provide educators with an instructional starting point by describing the skills and concepts that are most ready to be introduced, developed, or reinforced along a continuum of learning. This process is designed to assist classroom teachers, in particular, in translating the data from MAP assessment results into verbiage which is tightly aligned to the curriculum and facilitates the process of identifying student needs, whether those needs are around acceleration for high performing students or remediation for struggling students. Through the Learning Continuum reports, Test and Class Views supply global and student-specific information for tailoring instruction in which RIT scores are connected to skills and concepts students are ready to learn, helping to identify learning goals and targets to be shared with students and parents, as well as support efforts to create more personalized lesson plans. |
| **Performance Assessment:** | NWEA assessments currently include dichotomously scored items ranging from traditional multiple-choice items to technology enhanced items (TEI) that provide students with more interactive means to construct responses.<br><br>The different item types are selected based on analysis of the standards and take into |

| | |
|---|---|
| | consideration Bloom's cognitive process dimension and Webb's Depth of Knowledge (DOK) level(s) that can be targeted for any given standard. Our item innovation is driven by the content to be assessed – the need to provide authentic, engaging tasks that assess complex skills. To do this well, we look at the domain to be assessed and also at what is developmentally appropriate for children who see the test items. We are committed to continue developing new item types and scoring methodologies grounded in solid research and design.<br><br>MAP assessments are computer scored and feedback on student performance is provided within twenty-four hours, allowing teachers to make real time decisions about their classrooms. While we recognize the value in performance tasks, one of the purposes of interim assessments is to track student achievement and growth over time in order to predict performance on the summative assessment. At this time, we believe that performance tasks are best offered as culminating activities in summative tests.<br><br>The testing platform on which MAP resides provides a strong foundation for providing more engaging and "authentic" test items. As new items are added as field test items, the current practice of calibrating them can be extended to accommodate new item types, sets (clusters), and formats – placing all items on a single common content area scale. The interval characteristic of this scale allows achievement status within a content area to be tracked from one test occasion to another (i.e., measure growth). These changes in status (growth) can be modeled for groups of students, particularly for groups of students who are nested within classrooms/grade which is nested within a school. The 2015 RIT scale norms, provide a sophisticated example of how this can be done. However, simple growth estimates can also be obtained through basic arithmetic. In these cases, the norms can be used to determine where the growth estimates lie within a broad national context. |

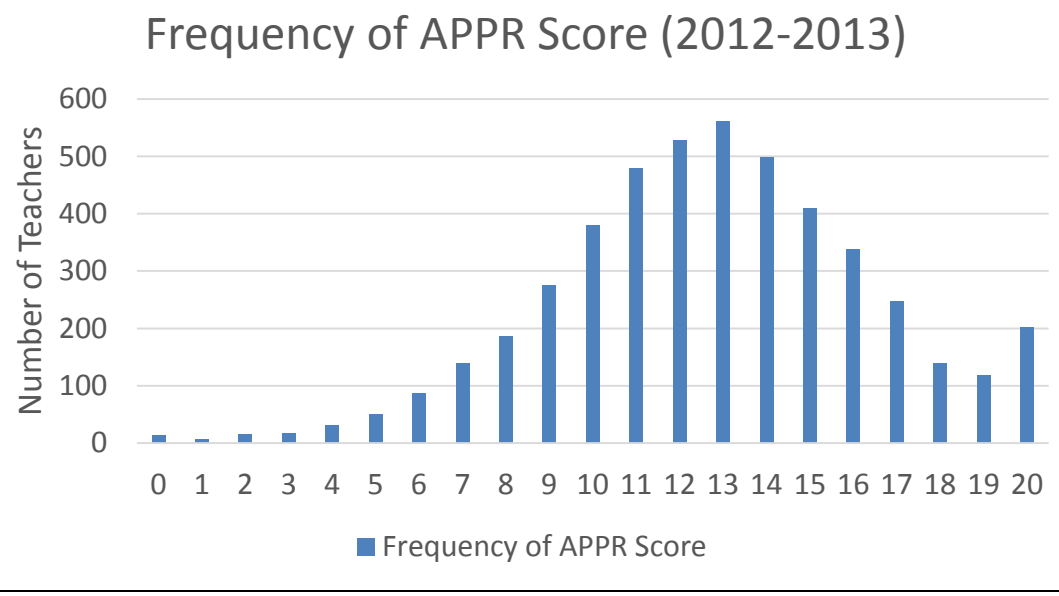
| | |
|---|---|
| | Because MAP assessments are administered in CAT and supported by high quality item pools anchored to vertical scales, the following measurement advantages can be achieved, which will substantially facilitate defensible judgments about educator effectiveness.<br>▪ **Broader Spectrum of Measurement.** Tests can provide scores with similar precision across the achievement range. This lies in contrast to fixed-form tests, in which students in the middle of a range are measured more reliably than students at the lower and higher extremes.<br>▪ **Precise Estimates.** Tests can provide superior precision over fixed-form tests used to estimate growth. The added precision affords more reliable estimates of student growth. When there is interest in referencing student or school level growth to national norms, the 2015 RIT scale norms allow comparisons to be based conditioned on weeks of instruction within a grade level, as well as on the starting score of the student (or school grade level). |
| **Efficient Time-Saving Assessments:** | Assessments may be administered in a variety of ways, either individually or in small or large groups, as long as administration occurs within a designated assessment window. Group administration conserves valuable instructional time and teacher resources and has no adverse effect on validity or reliability. Once testing is complete, results are available immediately in reports that demonstrate student performance at the individual, classroom, school, and district levels and allow real-time adjustment of instruction based on data to support personalized and highly impactful teaching and learning.<br><br>Using adaptive assessments to measure student achievement has a series of unique benefits, including: |

| | |
|---|---|
| | ▪ Increased Efficiency: Since test segments are allowed to conclude when estimates of student achievement are sufficiently precise, computer adaptive tests tend to be shorter than fixed form tests while yielding more instructionally valuable information about student achievement.<br>▪ Enhanced Precision: Adaptive tests are capable of enhancing the precision of student achievement estimates across the scale because, unlike fixed form tests with a single target information function for each form, adaptive tests offer different items closely tied to the student's achievement on previous questions.<br>▪ Improved Security: Because each student sees a unique test, increasing one's score by copying from a neighbor's test is virtually impossible.<br>▪ Reporting in Real-Time: The computer delivery of the assessment allows for immediate reporting of individual results, enabling educators to make better use of feedback from the assessment by making it easier to immediately use the results to inform instruction in real time.<br><br>On average, MAP tests take forty to sixty minutes to administer per domain. However, with assessments comprised of approximately four to six individual content goal areas, this amounts to less than ten minutes per goal area assessed.<br><br>Each test presents a student with a balanced number of items from each of the goals in order to gauge a student's performance level as it relates to key aspects of an academic area. With each MAP test administration, students receive an overall score for an academic area (e.g., reading or mathematics) as a whole as well as a score for each goal. For example, a student taking the MAP for Reading grades K – 2 assessment will receive a comprehensive score for reading as well as |

| | |
|---|---|
| | individual scores for each of the following goals:<br>▪ Foundational Skills (which includes assessment of alphabetic principle, phonics, and phonemic and phonological awareness skills);<br>▪ Language and Writing;<br>▪ Literature and Informational Texts (which includes assessment of a range of reading comprehension skills); and<br>▪ Vocabulary Use and Functions.<br><br>Thus, a single administration of a MAP assessment provides an efficient way to obtain a comprehensive picture of a student's performance in each of several key components of an academic area. In addition, MAP assessments provide considerable flexibility in the administration of the assessments. To accommodate a need for shorter administration times, the assessments can be paused at any time and resumed within a fourteen day period without impacting the test event. |
| **Technology:** | MAP assessments for grades K – 2 are delivered in the form of computerized adaptive tests (CAT) which utilize this technology to tailor item selection to the ability levels of examinees. All students take a unique version of the test, calibrated to a difficulty level where they will achieve approximately fifty percent correct answers. As a result, struggling students who typically become frustrated during testing, and high achievers who may find traditional tests boring, encounter a test that is appropriately challenging, which increases engagement and reduces the propensity to guess answers at random.<br><br>By creating a unique test for each student, educators receive highly accurate information about their students. Rather than simply indicating what a student might be able to do relative to grade-level standards, MAP tests indicate what a student is ready to learn relative to the applicable standards – not bound by grade. |

| | |
|---|---|
| | MAP assessments for students K – 2 meet the unique needs of early learners by displaying interactive elements and providing audio to ensure beginning readers understand the tasks presented in the assessment questions. Many items on the MAP for Primary Grades tests are interactive in nature, meaning students can manipulate and construct answers based on the learning being assessed.<br><br>Benefits of the adaptive technology of the MAP assessment system include:<br>▪ **Precise Data Faster:** Adaptive tests provide more accuracy in determining each student's achievement level using fewer items than a traditional fixed-form test, leading to shorter testing time. Upon completion of a MAP test, the assessment software calculates each student's score and immediately displays the score for the subject and goal areas via the end-of-test screen.<br>▪ **Flexible Reporting:** NWEA offers a robust suite of reports at the student, class, school, and district level in the MARC. The assessment software calculates each student's score and displays an overall RIT score via the end-of-test screen. MAP system reports and instructional resources are student-centric, research-based, and data-driven. The reports also provide data needed to inform instruction, evaluate programs, justify budget decisions, and help educators make key decisions.<br>▪ **Increased Student Confidence**: With adaptive testing, students gain confidence as they demonstrate what they are capable of doing without being bound by the restrictions inherent to a fixed-grade level instrument.<br>▪ **Broader Spectrum of Measurement:** Tests adapt to each student's instructional level independent of grade level, providing a greater depth of performance analysis. |

| | |
|---|---|
| | ▪ **Improved Security:** Each student develops a unique version of the test based on their performance, thereby reducing the likelihood that students may observe and use the answer of another student. |
| **Degree to which the growth model must differentiate across New York State's four levels of teacher effectiveness (only applicable to supplemental assessments):** | EdAnalytics can compute a standardized value-added measure by dividing each teacher's value-added score by the standard deviation of value added. The EdAnalytics model is designed specifically around understanding how teachers are differentiated in their impact on student growth. Each model is tested to make sure it can reliably detect differences in teacher impact and the overall scores are spread across the APPR scale. The graph below represents the distribution of APPR scores from the NWEA/EdAnalytics MAP APPR Growth work from 2012-2013 and shows clear distinctions between the Ineffective (0-2), Developing (3-8), Effective (9-17), and Highly Effective (18-20) categories. This graph is provided to illustrate the EdAnalytics model's ability to differentiate between categories rather than indicate the new distribution. It is expected that our proposed translation table will result in similar levels of differentiation. |



**STUDENT ASSESSMENTS FOR TEACHER AND PRINCIPAL EVALUATION**

**FORM *G***

# ATTESTATION OF TECHNICAL CRITERIA – SUPPLEMENTAL ASSESSMENTS WITH CORRESPONDING GROWTH MODELS

Please read each of the items below and check the corresponding box to ensure the fulfillment of the technical criteria outlined in the Technical Application on "FORM B-2".

PLEASE SUBMIT ONE "FORM G" FOR EACH APPLICANT. CO-APPLICANTS SHOULD SUBMIT SEPARATE FORMS.

**COMPLETE THIS SECTION:**

| **2.2(A) Narrative Overview of Proposed Supplemental Assessment and Associated Growth Model** | |
|---|---|
| This application contains a short overview of the assessment being proposed, including the intended purpose of the assessment, and how the assessment is administered. | ☒ |
| For supplemental assessments, this application contains a description of the growth model and how it is used in conjunction with the assessment. | ☐ ☒N/A |
| For K-2 assessments, this application contains evidence that the proposed assessment is consistent with this RFQ's requirement that the assessment not be a "Traditional Standardized Assessment" as defined above in the section "Definitions of Key Terms Used in this RFQ." | ☒ ☐ N/A |
| **2.2(B) Evidence of Capability** | |
| This application provides an overview of services provided by the Assessment Provider, including a description of the range of support / technical assistance that the Assessment Provider would provide to an LEA if selected by an LEA for this service. | ☒ |
| This application contains information as to whether the Applicant or Assessment Provider has been denied approval as a provider of assessment services in another state(s) and the reason(s) for such denial. If denied within New York State, the location and reason are indicated. | ☒ ☐ N/A |
| **2.2(C): Evidence of Copyright Owner/Assessment Representative History of Assessment Development** | |
| This application contains evidence that the Copyright Owner/Assessment Representative has a history of developing assessments of student learning (achievement or growth) for the purpose of making defensible judgments about educator effectiveness. | ☒ ☐ N/A |

| **2.2(D)-i: Technical Documentation Related to Assessment and Student Growth Score Properties: RELIABILITY**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
|---|---|
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for reliability:<br>• Student test scores have adequate levels of reliability (e.g., coefficient alpha > 0.75).<br><br>This application contains evidence of the *desired* criteria for reliability:<br>• Standard errors provided for students growth scores.<br>• Student growth classifications have adequate decision consistency.<br>• Teacher effectiveness classifications demonstrate adequate consistency.<br>*Examples include agreement statistics (e.g., kappa coefficients) based on simulation studies.* | **Check all that apply:**<br><br>☒<br><br>☒<br>☐<br>☐ |
| **2.2(D)-ii: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – ALIGNMENT**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for alignment validity:<br>• Evidence that test content is sufficiently aligned with New York State Learning Standards and covers a range of measurable standards. Documentation that demonstrates that:<br>(a) at least 80% of the test measures content aligned with NYS learning standards,<br>(b) no more than 20% of test content is aligned with other learning standards or objectives, and<br>(c) a range of content from the NYS learning standards is measured<br><br>*Note: Other relevant standards can be proposed if NYS Learning Standards do not apply to subject area.*<br><br>This application contains evidence of the *desired* criteria for alignment validity:<br>• 100% alignment between NYS Learning Standards and assessment. | **Check all that apply:**<br><br>☒<br><br>☒ |
| **2.2(D)-iii: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – RELATIONS TO OTHER VARIABLES**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for validity in relation to other variables:<br>• Evidence students' growth scores are correlated with other measures of student progress (e.g., r > .5 with measures such as the number of objectives | **Check all that apply:** |

| | |
|---|---|
| mastered by a student over the course of the year, teachers' ratings of students' progress, or scores from other assessments).<br><br>This application contains evidence of the *desired* criteria for validity in relation to other variables:<br>• Evidence teacher effectiveness ratings are positively correlated (e.g., r > .5) with other measures of teaching effectiveness. | ☒<br><br>☐ |
| **2.2(D)-iv: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – INTERNAL STRUCTURE**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for validity of internal structure:<br>• Scale properties appropriate for growth model used (*see notes*). Total scores and subscores on student assessments should be supported by dimensionality analyses (e.g., IRT residual analyses, factor analyses).<br><br>This application contains evidence of the *desired* criteria for validity of internal structure:<br>• Evidence students' scores are on an interval scale.<br><br>**Notes: If **gain score model** is used, evidence is needed that students' pretest and posttest scores are on the same scale. If **student growth percentile** model used, justification for the number of years included in the model should be provided. If **growth-to-proficiency**, projection, or **value-added models** are used, evidence is needed that the model explains a significant amount of variability in student achievement. Also, models should demonstrate robustness to missing data.* | **Check all that apply:**<br><br>☒<br><br>☒ |
| **2.2(D)-v: Technical Documentation Related to Assessment and Student Growth Score Properties: UTILITY AND COMPREHENSIBILITY**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for utility and comprehensibility:<br>• Technical documentation that describes how student growth and educator effectiveness are calculated.<br><br>This application contains evidence of the *desired* criteria for utility and comprehensibility:<br>• Student growth reports support instructional improvement. Resources and supporting materials available to the field. | **Check all that apply:**<br><br>☒<br><br>☒ |
| **2.2(E)-i: Technical Documentation Related to Aggregating Student-Level Growth Scores to Teacher-Level Scores: CREATION OF TEACHER LEVEL SCORES** | |

| | |
|---|---|
| **For supplemental assessments used in conjunction with growth models:**<br>This application includes a narrative description of how student-level scores are aggregated to create a single teacher-level score for each teacher. | ☒ ☐ N/A |
| **2.2(E)-ii: Technical Documentation Related to Aggregating Student-Level Growth Scores to Teacher-Level Scores: EXCLUSION RULES** | |
| This application includes a description of any exclusion rules that remove students associated with a given teacher from the teacher's teacher-level score (either through a growth model or in conjunction with an SLO). | ☒ ☐ N/A |
| **2.2(F): Technical Documentation Related to Converting Teacher-Level Growth Score to New York State's 0-20 APPR Scale** | |
| This application includes a crosswalk that maps scores on the assessment's aggregated teacher-level growth score to the required New York State teacher and principal evaluation metric, which ranges from 0-20.<br><br>This application includes procedures for converting teacher-level growth scores to the 0-20 APPR scale comply with the New York Standards for each evaluation rating category, which are based on the following definitions.<br><br>**For supplemental assessments used in conjunction with growth models:**<br>This application includes an explanation of the assignment of HEDI rating categories based on the following ranges:<br>• Highly Effective: results are well-above State average* for similar students<br>• Effective: results meet State average* for similar students<br>• Developing: results are below State average* for similar students<br>• Ineffective: Results are well-below State average* for similar students | ☒<br><br>☒<br><br>☒ ☐ N/A |
| **2.2(G)-i: Technical Documentation Related to Fairness: TEST TAKERS**<br>Consistent with the new Testing Standards (2014), there is an increased focus in the industry on fairness of assessments and their uses. Please provide evidence of fairness for both the proposed assessment and, if applicable, the proposed growth model. | |
| This application includes evidence that the proposed assessments are fair to all test takers (e.g., Differential Item Functioning [DIF] / bias information, fairness evaluation / sensitivity review plan.) | ☒ |
| **2.2(G)-ii: Technical Documentation Related to Fairness: TEACHER GROWTH SCORES** | |
| This application includes evidence of fairness of the proposed aggregated teacher growth scores (e.g., lack of correlation between aggregated teacher growth scores and student demographics).<br><br>The evidence of fairness of the proposed aggregated teacher growth scores includes an explanation of how the growth model incorporates (a) prior academic history, (b) poverty, (c) students with disabilities, and (d) English language learners. | ☒<br><br>☒ ☐ N/A |

**To be completed by the Copyright Owner/Assessment Representative of the assessment being proposed and, where necessary, the co-applicant LEA:**

| | |
|---|---|
| The Northwest Evaluation Association (NWEA)<br>1. Name of Organization (PLEASE PRINT/TYPE) | 4. Signature of Authorized Representative (PLEASE USE **BLUE** INK) |
| Geri Cohen<br>2. Name of Authorized Representative (PLEASE PRINT/TYPE) | 09/30/2015<br>5. Date Signed |
| Vice President of Finance & CFO<br>3. Title of Authorized Representative (PLEASE PRINT/TYPE) | |

| | |
|---|---|
| NA<br>1. Name of LEA (PLEASE PRINT/TYPE) | 4. Signature of School Representative (PLEASE USE **BLUE** INK) |
| 2. School Representative's Name (PLEASE PRINT/TYPE) | 5. Date Signed |
| 3. Title of School Representative (PLEASE PRINT/TYPE) | |

**STUDENT ASSESSMENTS AND ASSOCIATED GROWTH MODELS FOR TEACHER AND PRINCIPAL EVALUATION**

**FORM *C***

# PUBLICLY AVAILABLE SERVICES SUMMARY

This form will be posted on the New York State Education Department's Web site and distributed through other means for all applications that are approved in conjunction with this RFQ to allow districts and BOCES to understand proposed offerings in advance of directly contacting Assessment Providers regarding potential further procurements.

| **Assessment Provider Information** | |
|---|---|
| Name of Assessment Provider: | The Northwest Evaluation Association (NWEA) |
| Assessment Provider Contact Information: | Michelle LaPlatney, NWEA Account Executive<br>Ph: (973) 769-9985<br>Email: michelle.laplatney@nwea.org<br>121 NW Everett Street<br>Portland, OR 97209 |
| Name of Assessment: | Measures of Academic Progress (MAP) |
| Nature of Assessment: | ☐ ASSESSMENT FOR USE WITH STUDENT LEARNING OBJECTIVES WITH A TARGET SETTING MODEL; OR<br><br>☒ SUPPLEMENTAL ASSESSMENT WITH AN ASSOCIATED GROWTH MODEL:<br>☐ GAIN SCORE MODEL<br>☐ GROWTH-TO-PROFICIENCY MODEL<br>☐ STUDENT GROWTH PERCENTILES<br>☐ PROJECTION MODELS<br>☒ VALUE-ADDED MODELS<br>☐ OTHER: |
| What are the grade(s) for which the assessment can be used to generate a 0-20 APPR score? | Grades 2 – 10 |
| What are the subject area(s) for which the assessment can be used to generate a 0-20 APPR score? | Mathematics, Reading, and Language Usage |
| What are the technology requirements associated with the assessment? | All NWEA reports for MAP assessments are available online through the MAP Administration and Reporting Center (MARC), which provides a comprehensive set of intuitive web-based reports, instructional content, data tools, and comparative information sources. The center is accessible to educators and administrators from any location with an Internet connection, at any time outside of scheduled maintenance. Maintenance typically occurs once per month on the weekends, and NWEA provides an advance schedule to partners. |

| | |
|---|---|
| | For current technical requirements, please see: https://teach.mapnwea.org/impl/QRM2_System_Requirements_QuickRef.pdf<br><br>MAP is also supported for current partners who have been using our client server platform. Technical requirements for this platform are unchanged from our previous application. However, all new partners will implement MAP assessments via the web-based platform described above. |
| Is the assessment available, either for free or through purchase, to other districts or BOCES in New York State? | ☒ YES<br><br>☐ NO |

**Please provide an overview of the assessment for districts and BOCES. Please include:**

- **A description of the assessment;**
- **A description of how the assessment is administered;**
- **A description of how scores are reported (include links to sample reports as appropriate);**
- **A description of how the Assessment Provider supports implementation of the assessment, including any technical assistance. (3 pages max)**

## Overview of MAP Assessments

The Northwest Evaluation Association™ (NWEA™) is pleased to provide this introduction to the Measures of Academic Progress® (MAP®) assessment system. Our web-based, item-level adaptive interim assessments measure student achievement in mathematics, reading, and language usage in grades 2 – 10. We offer our experience, expertise, and research-based assessments to districts and BOCES in New York to empower educators to accurately measure student achievement and growth within the school year and across years. The precise data MAP assessments provide, aligned to New York State P – 12 Common Core Learning Standards (NYSCCLS), will allow educators to make the kinds of immediate instructional decisions that can affect positive change for every student.

MAP assessments accurately reflect each student's instructional level and provide educators immediate, highly accurate, detailed data about what students know and what they are ready to learn next. MAP assessments are recognized as one of the highest quality, research-based interim assessment systems available due to the strength of the vertical MAP scale, our test reliability and validity, adaptive test algorithms, large item pools, and the stringent item development processes we follow. Our system of support includes robust professional development services, account management services, and technical support to ensure educators can administer the tests easily, understand the results, and take action to improve instruction.

As one of the founders of the adaptive testing movement, NWEA has over thirty years of experience in this field, and well understands the value of providing rich and efficient testing experiences that are tailored to the individual learning of each student.

MAP assessments are recognized as one of the most accurate measures of student achievement and growth in the market, and are used by over 7,400 NWEA partners including state departments of education, school districts, private schools, charter schools, foundations, universities, school reform groups, the Bureau of Indian Education (BIE), and international schools.

With the MAP assessment system, educators also gain:

- A stable, grade-independent, vertical scale that measures growth with precision even as standards and education continue to change
- Powerful reporting options for multiple stakeholders, including district leadership, principals, teachers, students, and parents
- Fast access to data, as MAP produces student scores immediately after test completion
- Longitudinal data to track student growth over time
- Growth and achievement norms that allow for the national comparison of the achievement and growth patterns of students in your district or BOCES with students in all fifty states
- Experienced implementation, technical support, and account management personnel to ensure smooth onboarding and administration of assessments
- Secure and reliable data available for export to student information systems and data warehouses
- A stable testing platform that can reliably deliver assessments and results, even in large districts. The platform supports 120,000 concurrent test events with item response times within milliseconds

## MAP Assessment Administration

We administer more than forty million MAP assessments annually throughout the world. Our assessments have been used to target and improve classroom instruction and advance student achievement across the state, across the nation, and across the globe.

MAP assessments are adaptive, meaning that each test is designed to optimally engage each student by adjusting to his or her instructional level, at the item level, through our adaptive test engine. The assessment begins by delivering the student a grade-level question. If the student answers the question correctly, the test taker is rewarded with a more difficult question. Conversely, an incorrect response triggers the delivery of an easier second question.

All students take a unique version of the test, calibrated to a difficulty level where they will achieve approximately fifty percent correct answers. As a result, struggling students who typically become frustrated during testing, and high achievers who may find traditional tests boring, encounter a test that is appropriately challenging, which increases engagement and reduces the propensity to guess answers at random.

By creating a unique test for each student, educators receive highly accurate information about their students. Rather than simply indicating what a student might be able to do relative to grade-level standards, MAP tests indicate what a student is ready to learn relative to the NYSCCLS – not bound by grade.

Our partners' experience of the MAP assessment system is that it is intuitive, easy to operate, and engaging for both students and educators. The MAP system has a visually appealing interface that is simple to navigate, providing users with embedded, page-specific online help, guides, and tutorials on-demand. All of this supports our ability to keep educators' time invested in learning the application to a minimum, while maximizing the ability to obtain useful and actionable information from the data.

## Fast, User-Friendly Reporting

Upon completion of a MAP test, the assessment software calculates each student's score and immediately displays the score for the subject and goal areas via the end-of-test screen. While student reports are available immediately after testing, reports that aggregate data at the classroom, school, and district level become available after testing for that group of students is complete. All reports are available online using the MAP Administration and Reporting Center (MARC) and can be accessed from any location with an Internet connection. A review of the reports available in the MARC can be found here: https://www.nwea.org/assessments/map/reporting-data/featured-reports-measures-academic-progress-map-map-primary-grades/.

For more information on the reporting offered by the MAP assessment suite, please visit our reporting resources page with Quick Facts, a short video overview of standard MAP assessment reporting, and additional information on customized reports at: https://www.nwea.org/assessments/map/reporting-data/.

## Support Services

MAP assessments are a fully hosted solution and are easy to implement and to administer. We also provide outstanding support to partners throughout our relationships to make sure we are effectively meeting their needs. The NWEA Partner Accounts and Partner Services teams are responsible for providing our partners with account management, professional development, implementation, and technical support. A culture of continuous improvement coupled with hiring practices focused on

selecting individuals with strong customer-service orientation has created an excellent team that is well respected by our partners. Our staff is well-versed in the benefits and challenges inherent in the implementation and ongoing delivery of computer-based adaptive assessment systems.

**Implementation Services**

During a partner's first testing season, NWEA assigns an Implementation Support Specialist to proactively guide designated 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.

After a partner has completed their initial implementation of the MAP system, ongoing account management duties are transitioned from the Implementation Support Specialist to the Account Manager.

**Account Management**

The designated Account Manager is the point of contact at NWEA for any partner questions. Our Account Managers work out of our national headquarters in Portland, Oregon or out of their remote home offices located around the country, and are available via phone or email on a daily basis. Most inquiries receive a response within twenty-four to forty-eight hours. Account Managers schedule periodic check-in meetings with partners to answer questions, follow up on any open issues, schedule additional professional development as needed, and collaborate with staff on their plans for the next season of testing.

**Technical Support**

Our highly skilled Technical Support Team is available by toll-free Support line, email, or chat Monday through Friday, 7:00 a.m. to 8:00 p.m. Eastern Time (EST), excluding federally recognized holidays observed by NWEA. Our Technical Support staff is well-versed in the implementation of our assessment systems and can provide assistance with generating roster files, configuring system components, accessing online reports, and answering any questions that arise in the use of the systems.

***Reciprocal Partner Communications***

Incoming calls, emails, and chats are routed through a tiered support system for effective triaging based on the nature and urgency of the question or issue. Our call routing and escalation processes efficiently route service requests to the appropriate personnel. NWEA staff respond with accurate, timely, courteous, and consistent service. To optimize responsiveness and maintain consistently high customer satisfaction, we use established Information Technology Infrastructure Library-based (ITIL) escalation protocols.

Technical Support staff log all partner inquiries and the resolution in a Customer Relationship Management (CRM) application, providing an historic record of all partner support interactions. The management team regularly reviews these log reports to identify trends, escalate bug fixes, and analyze options for enhancements to be included in future product releases.

NWEA also provides product release notes for our partners to communicate new product features and other "under the hood" changes that improve system performance and stability. These monthly Partner Update newsletters are hyperlinked from the MARC, and also provide a synopsis of upcoming partner-facing changes that explain what changes or updates will be taking place in the coming days. Prior to the release, partners receive a partner update email that highlights important partner-facing changes.

**Please provide an overview of the student-level growth model or target setting model for SLOs for districts and BOCES, along with how student-level growth scores are aggregated to the create teacher-level scores, and how those teacher-level scores are converted to New York State's 0-20 metric.**

NWEA is partnering with researchers at Education Analytics Inc. (EdAnalytics) to provide districts in New York with value-added estimates of educator effectiveness. The use of this approach is the best means to ensure that student results accurately capture the contributions educators made to the growth of their students, by eliminating noise from factors external to the educator and outside his or her control that could lead to mistaken generalizations about that individual's performance.

EdAnalytics has an established value-added process that yields estimates of the contribution of educational units (classrooms and schools) and agents (teachers and principals) to student achievement or other student outcomes, after controlling for external (non-school) sources of student achievement growth, including prior student achievement and student and family characteristics. EdAnalytics' objective is to facilitate valid and fair comparisons of productivity with respect to student outcomes, given that schools often serve very different student populations.

EdAnalytics collects data files from a variety of sources including the NWEA test files, districts' SIRS data files, and templates especially created for this work. By combining all these files, EdAnalytics can gain a comprehensive view of both school and non-school factors that may have affected the growth of a particular teacher's students. EdAnalytics' methodology works by comparing the growth of each student to observationally similar students (students with similar starting points and characteristics). If students in a teacher's classroom tend to grow faster than observationally similar students, the teacher receives a high value-added rating, meaning the value that teacher added to student growth was greater than expected.

EdAnalytics produces results for each teacher that indicates how a teacher's students grew relative to the average of similar students. These results are then translated to the proposed 0-20 HEDI scale. In EdAnalytics' proposed system, a teacher whose students grew at a typical rate for similar students will receive a "sixteen", the middle of the "Effective" range. Results below sixteen do not indicate that students lost knowledge, but rather indicate that student growth was below-average as compared to the growth of similar students.

In order to create and update its methodology throughout the course of this existing work, EdAnalytics has employed its "co-build" approach where stakeholders from participating districts participate in New York MAP Value-Added Advisory Council meetings. At these meetings, stakeholders from participating districts along with experts from EdAnalytics and NWEA evaluate the ongoing development of the growth metrics and ensure the methodology used continues to address evolving policy needs while remaining compliant with state regulations.

| **New York State Next Generation Assessment Priorities**<br>Please provide detail on how the proposed supplemental assessment l or assessment to be used with SLOs addresses each of the Next Generation Assessment Priorities below. | |
|---|---|
| **Characteristics of Good ELA and Math Assessments (only applicable to ELA and math assessments):** | The MAP assessments are consistent with many of the criteria supplied by the Achieve the Core Assessment Evaluation Tool (AET). However, the adaptive and across-grade nature of the MAP assessments mean that some of AET criteria do not apply since those criteria focus on within grade assessments. The MAP assessments are specifically designed to cross grades as this structure allows the assessments to measure where each individual student is performing, show growth, and provide teachers with more precise information about what their students know.<br><br>The AET criteria for ELA tests in grades 3 – 12 include: the importance of the complexity and quality of the texts, test questions that are standards-based, texts that reflect the distribution of text types and genres required by the state standards, vocabulary items that assess words in context and focus on central ideas in the text, items around conventions and writing strategies that focus on the standards and actual practice, a variety of items types that are appropriate to the standards, and test blueprints that reflect the standards. The AET for ELA tests does not specifically mention K – 2, however MAP assessments in the K – 10 grade range adhere to the information provided below.<br><br>The items used in the MAP assessment item pools are all aligned to the NYSCCLS and the test blueprints (goal structures) are built to reflect the organization of the NYSCCLS and provide evidence of alignment to the standards. For example, MAP for Reading assessments available for grades K – 10 cover Vocabulary, Literary Texts, and Informational Texts explicitly; the MAP Language Usage tests available for grades 2 – 10 assess writing conventions such as spelling, mechanics, grammar, and usage as well as selected writing strategies related to planning, composing, revising, and editing in context.<br><br>Recognizing that some Standards do not lend themselves to selected-response items, NWEA continues to add alternate item types such as:<br>▪ Technology-enhanced items: interactive items that students can manipulate to construct answers based on the learning assessed<br>▪ Drag and Drop: a student drags one or more objects (e.g., numbers, words, pictures) from one location of the screen to another to create an answer |

| | |
|---|---|
| | ▪ Click and Pop: a student clicks on one or more answer objects (e.g., numbers, words, pictures) that automatically move to a pre-selected location on the screen<br>▪ Hot Spot: a student clicks on one or more objects (e.g., geometric shapes, text, symbols) to select answer option(s); items indicated by the student are highlighted<br>▪ Common stimulus reading items**:** sets of items associated with a single literary or nonfiction stimulus (e.g. an extended passage). Students read the passage and answer a series of selected-response items that target a variety of skills, requiring students to engage different cognitive processes. Extended passages allow for a more authentic and sustained reading experience, where students can respond holistically to a complete text of appropriate rigor<br>▪ Items containing video and animation<br>▪ Items that use drawing and orienting lines and figures<br>▪ Virtual performance tasks<br>▪ Constructed-response items<br><br>All NWEA items go through a rigorous item development and review process. The process yields items with strong alignments to the breadth and depth of the NYSCCLS. In order to achieve this, we have developed a deep understanding of the standards and use a variety of approaches and item types to assess them. Items are developed and reviewed through a variety of lenses, including how they align to the targeted standard and grade level, how they adhere to the principles of Universal Design, and whether they are free from potential bias and sensitivity issues. Additionally, the literary and informational texts used by items are evaluated both quantitatively and qualitatively, looking at criteria such as knowledge demand, language conventionality, and clarity in order to determine the readability and complexity of each passage.<br><br>The AET criteria for Math tests in grade K – 12 include some criteria that assume an on-grade-level, fixed-form assessment: a focus on the concept of Major Work for the grade, items should not assess topics before they are introduced in the standards, reflect the grade-by-grade progression of the standards, and score reporting that reflects the emphasis of the grade. NWEA assessments are adaptive tests that are designed to assess students where they are regardless of grade level and to show growth regardless of grade level. Students will see items that are aligned to standards above or below their grade level. However, NWEA's RIT scores and Learning Continuum reports make it simple to determine where students are performing with regard to grade level |

| | |
|---|---|
| | standards. They also make it possible to see what skills and concepts students who are performing above or below grade level know. The MAP assessments for mathematics do contain items that align to the standards and therefore reflect the three aspects of rigor and the connection between content standards and practice standards.<br><br>In addition to specifics about the ELA and Mathematics assessments, AET provides guidance about indicators of a quality assessment. As mentioned above, all items go through a multi-stage item development and review process to ensure high quality items. NWEA assessments are built to be student worthy assessments. The purpose of the tests, the data and reporting needs, and learning targets are all considered as part of the test development process. Items are then aligned to the learning targets and go through a field testing and calibration process to place them on the measurement scale and to ensure accuracy and validity. Test designs are carefully constructed to include the relevant content, technical, and psychometric information needed for test construction. The tests and items that result from these defined processes offer valid content, reliability in terms of valid data for students at all levels and across years, and fairness across student populations.<br><br>The items presented to a student in any given test event are determined by the individual student's achievement level and by the test's goal structure. These goal structures group all assessable standards into goal areas that represent content domains and sub-goals that represent common groupings of grade level expectations, which cover related topics along the learning continuum within each standard. Each student is administered a balanced number of items in each goal area to provide an overall score for the content area (mathematics, reading, or language usage), as well as goal area scores.<br><br>Because MAP tests are adaptive and designed to provide data about students across the achievement continuum – including students who are performing below level or above level – the item pools that support these tests are very large and include items that may range in complexity from the most basic "building block" aspect of a skill to analytical or evaluative aspects of the skill. MAP assessments are designed to assess students where they are, regardless of grade level. For example, the MAP 2 – 5 mathematics test draws from an item pool containing items aligned to some of the standards below grade two and some above grade five. This way, if a student is performing below second grade or above fifth grade, the test can identify those specific skills and concepts. However, the MAP assessment for mathematics does contain items that align to the standards and therefore reflect the three aspects |

| | |
|---|---|
| | of rigor and the connection between content standards and practice standards.<br><br>In addition to specifics about the ELA and Mathematics assessments, AET provides guidance about indicators of a quality assessment. As mentioned above, all items go through a multi-stage item development and review process to ensure high quality items. NWEA Assessments are built to be student worthy assessments. The purpose of the tests, the data and reporting needs, and learning targets are all considered as part of the test development process. Items are then aligned to the learning targets, and go through a field testing and calibration process to place them on the measurement scale and to ensure accuracy and validity. Test designs are carefully constructed to include the relevant content, technical, and psychometric information needed for test construction. The tests and items that result from these defined processes offer valid content, reliability in terms of valid data for students at all levels and across years, and fairness across student populations. |
| **Assessments Woven Tightly Into the Curriculum:** | NWEA believes that each and every student matters, and we offer assessments designed to help guide meaningful classroom instruction. MAP assessments offer a personalized experience for students by adapting to each student's learning level – precisely measuring student progress and growth for each individual. Assessments are designed to be completed within a short amount of time (forty to sixty minutes per domain) and to provide teachers with robust information within twenty-four hours about what each student knows and is ready to learn, which can be used to inform classroom instruction.<br><br>MAP assessments provide teachers with a means to measure the growth and progress of every student over time, regardless of whether a student is performing on, above, or below grade level. In addition, the assessments compare and predict student achievement and growth over time via NWEA achievement and growth norms. These data can be used by teachers to personalize instruction quickly for 1:1, small group, or whole class activities. Teachers can also use the data to support efforts to engage students in achieving personalized learning goals and progress via student and family goal setting activities.<br><br>MAP assessments include our proprietary interactive tool for teachers, the Learning Continuum. Teachers can use the Learning Continuum's information to streamline instructional planning, differentiate instruction for both individual students and skill-based activity groups, and better engage students in |

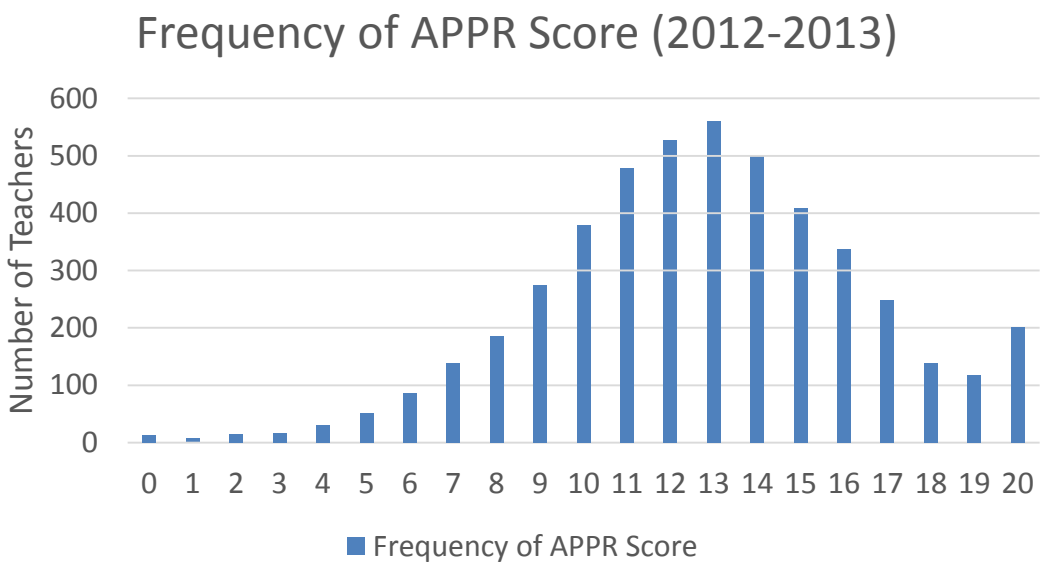
| | |
|---|---|
| | their learning. It is a powerful shortcut to understanding which skills students are ready to learn.<br><br>Within the Learning Continuum, learning statements provide educators with an instructional starting point by describing the skills and concepts that are most ready to be introduced, developed, or reinforced along a continuum of learning. This process is designed to assist classroom teachers, in particular, in translating the data from MAP assessment results into verbiage which is tightly aligned to the curriculum and facilitates the process of identifying student needs, whether those needs are around acceleration for high performing students or remediation for struggling students. Through the Learning Continuum reports, Test and Class Views supply global and student-specific information for tailoring instruction in which RIT scores are connected to skills and concepts students are ready to learn, helping to identify learning goals and targets to be shared with students and parents, as well as support efforts to create more personalized lesson plans. |
| **Performance Assessment:** | NWEA assessments currently include dichotomously scored items ranging from traditional multiple-choice items to technology enhanced items (TEI) that provide students with more interactive means to construct responses. MAP for Reading tests also include reading item sets that use an extended passage as a stimulus for multiple items. The extended passage allows for a more authentic and sustained reading experience in which students can respond holistically to a text of appropriate complexity and rigor. Such passages allow for closer reading and support items related to concepts such as citing and analyzing evidence, evaluating arguments and claims, text organization and structure, plot development, characterization, theme, and author's style.<br><br>The different item types are selected based on analysis of the standards and take into consideration the Bloom's cognitive process dimension and Webb's Depth of Knowledge (DOK) level(s) that can be targeted for any given standard. Our item innovation is driven by the content to be assessed – the need to provide authentic, engaging tasks that assess complex skills. To do this well, we look at the domain to be assessed and also at what is developmentally appropriate for children who see the test items. We are committed to continue developing new item types and scoring methodologies grounded in solid research and design.<br><br>MAP assessments are computer scored and feedback on student performance is provided within twenty-four hours, allowing teachers to make real time decisions about their classrooms. While we recognize the value in performance tasks, one of the purposes of interim assessments is to track |

| | |
|---|---|
| | student achievement and growth over time in order to predict performance on the summative assessment. At this time, we believe that performance tasks are best offered as culminating activities in summative tests.<br><br>The testing platform on which MAP resides provides a strong foundation for providing more engaging and "authentic" test items. As new items are added as field test items, the current practice of calibrating them can be extended to accommodate new item types, sets (clusters), and formats – placing all items on a single common content area scale. The interval characteristic of this scale allows achievement status within a content area to be tracked from one test occasion to another (i.e., measure growth). These changes in status (growth) can be modelled for groups of students, particularly for groups of students who are nested within classrooms/grade which is nested within a school. The 2015 RIT scale norms, provide a sophisticated example of how this can be done. However, simple growth estimates can also be obtained through basic arithmetic. In these cases, the norms can be used to determine where the growth estimates lie within a broad national context.<br><br>Because MAP assessments are administered in CAT and supported by high quality item pools anchored to vertical scales, the following measurement advantages can be achieved which will substantially facilitate defensible judgments about educator effectiveness.<br>▪ **Broader Spectrum of Measurement.** Tests can provide scores with similar precision across the achievement range. This lies in contrast to fixed-form tests, in which students in the middle of a range are measured more reliably than students at the lower and higher extremes.<br>▪ **Precise Estimates.** Tests can provide superior precision over fixed-form tests used to estimate growth. The added precision affords more reliable estimates of student growth. When there is interest in referencing student or school level growth to national norms, the 2015 RIT scale norms allow comparisons to be based conditioned on weeks of instruction within a grade level, as well as on the starting score of the student (or school grade level). |
| **Efficient Time-Saving Assessments:** | Assessments may be administered in a variety of ways, either individually or in small or large groups, as long as administration occurs within a designated assessment window. Group administration conserves valuable instructional time and teacher resources and has no adverse effect on validity or reliability. Once testing is complete, results are available immediately in reports that demonstrate student performance at the individual, classroom, school, and district levels and allow real-time adjustment of instruction |

| | |
|---|---|
| | based on data to support personalized and highly impactful teaching and learning.<br><br>Using adaptive assessments to measure student achievement has a series of unique benefits, including:<br>▪ Increased Efficiency: Since test segments are allowed to conclude when estimates of student achievement are sufficiently precise, computer adaptive tests tend to be shorter than fixed form tests while yielding more instructionally valuable information about student achievement.<br>▪ Enhanced Precision: Adaptive tests are capable of enhancing the precision of student achievement estimates across the scale because, unlike fixed form tests with a single target information function for each form, adaptive tests offer different items closely tied to the student's achievement on previous questions.<br>▪ Improved Security: Because each student sees a unique test, increasing one's score by copying from a neighbor's test is virtually impossible.<br>▪ Reporting in Real-Time: The computer delivery of the assessment allows for immediate reporting of individual results, enabling educators to make better use of feedback from the assessment by making it easier to immediately use the results to inform instruction in real time.<br><br>On average, MAP tests take forty to sixty minutes to administer per domain. However, with assessments comprised of approximately four to six individual content goal areas, this amounts to less than ten minutes per goal area assessed.<br><br>Each test presents a student with a balanced number of items from each of the goals in order to gauge a student's performance level as it relates to key aspects of an academic area. With each MAP test administration, students receive an overall score for an academic area (e.g., reading or mathematics) as a whole as well as a score for each goal.<br><br>Thus, a single administration of a MAP assessment provides an efficient way to obtain a comprehensive picture of a student's performance in each of several key components of an academic area. In addition, MAP assessments provide considerable flexibility in the administration of the assessments. To accommodate a need for shorter administration times, the assessments can be paused at any time and resumed within a fourteen day period without impacting the test event. |

| **Technology:** | MAP assessments are delivered in the form of computerized adaptive tests (CAT) which utilize this technology to tailor item selection to the ability levels of examinees. All students take a unique version of the test, calibrated to a difficulty level where they will achieve approximately fifty percent correct answers. As a result, struggling students who typically become frustrated during testing, and high achievers who may find traditional tests boring, encounter a test that is appropriately challenging, which increases engagement and reduces the propensity to guess answers at random.<br><br>By creating a unique test for each student, educators receive highly accurate information about their students. Rather than indicating what a student might be able to do relative to grade-level standards, MAP tests indicate what a student is ready to learn relative to the applicable standards – not bound by grade.<br><br>Benefits of the adaptive technology of the MAP assessment system include:<br>▪ **Precise Data Faster:** Adaptive tests provide more accuracy in determining each student's achievement level using fewer items than a traditional fixed-form test, leading to shorter testing time. Upon completion of a MAP test, the assessment software calculates each student's score and immediately displays the score for the subject and goal areas via the end-of-test screen.<br>▪ **Flexible Reporting:** NWEA offers a robust suite of reports at the student, class, school, and district level in the MARC. The assessment software calculates each student's score and displays an overall RIT score via the end-of-test screen. MAP system reports and instructional resources are student-centric, research-based, and data-driven. The reports also provide data needed to inform instruction, evaluate programs, justify budget decisions, and help educators make key decisions.<br>▪ **Increased Student Confidence**: With adaptive testing, students gain confidence as they demonstrate what they are capable of doing without being bound by the restrictions inherent to a fixed-grade level instrument.<br>▪ **Broader Spectrum of Measurement:** Tests adapt to each student's instructional level independent of grade level, providing a greater depth of performance analysis.<br>▪ **Improved Security:** Each student develops a unique version of the test based on their performance, thereby reducing the likelihood that students may observe and use the answer of another student. |
|---|---|

| | |
|---|---|
| **Degree to which the growth model must differentiate across New York State's four levels of teacher effectiveness (only applicable to supplemental assessments):** | EdAnalytics can compute a standardized value-added measure by dividing each teacher's value-added score by the standard deviation of value added. The EdAnalytics model is designed specifically around understanding how teachers are differentiated in their impact on student growth. Each model is tested to make sure it can reliably detect differences in teacher impact and the overall scores are spread across the APPR scale. The graph below represents the distribution of APPR scores from the NWEA/EdAnalytics MAP APPR Growth work from 2012-2013 and shows clear distinctions between the Ineffective (0-2), Developing (3-8), Effective (9-17), and Highly Effective (18-20) categories. This graph is provided to illustrate the EdAnalytics model's ability to differentiate between categories rather than indicate the new distribution. It is expected that our proposed translation table will result in similar levels of differentiation. |



**STUDENT ASSESSMENTS FOR TEACHER AND PRINCIPAL EVALUATION**

**FORM *G***

# ATTESTATION OF TECHNICAL CRITERIA – SUPPLEMENTAL ASSESSMENTS WITH CORRESPONDING GROWTH MODELS

Please read each of the items below and check the corresponding box to ensure the fulfillment of the technical criteria outlined in the Technical Application on "FORM B-2".

PLEASE SUBMIT ONE "FORM G" FOR EACH APPLICANT. CO-APPLICANTS SHOULD SUBMIT SEPARATE FORMS.

**COMPLETE THIS SECTION:**

| **2.2(A) Narrative Overview of Proposed Supplemental Assessment and Associated Growth Model** | |
|---|---|
| This application contains a short overview of the assessment being proposed, including the intended purpose of the assessment, and how the assessment is administered. | ☒ |
| For supplemental assessments, this application contains a description of the growth model and how it is used in conjunction with the assessment. | ☐ ☒N/A |
| For K-2 assessments, this application contains evidence that the proposed assessment is consistent with this RFQ's requirement that the assessment not be a "Traditional Standardized Assessment" as defined above in the section "Definitions of Key Terms Used in this RFQ." | ☐ ☒N/A |
| **2.2(B) Evidence of Capability** | |
| This application provides an overview of services provided by the Assessment Provider, including a description of the range of support / technical assistance that the Assessment Provider would provide to an LEA if selected by an LEA for this service. | ☒ |
| This application contains information as to whether the Applicant or Assessment Provider has been denied approval as a provider of assessment services in another state(s) and the reason(s) for such denial. If denied within New York State, the location and reason are indicated. | ☒ ☐ N/A |
| **2.2(C): Evidence of Copyright Owner/Assessment Representative History of Assessment Development** | |
| This application contains evidence that the Copyright Owner/Assessment Representative has a history of developing assessments of student learning (achievement or growth) for the purpose of making defensible judgments about educator effectiveness. | ☒ ☐ N/A |

| **2.2(D)-i: Technical Documentation Related to Assessment and Student Growth Score Properties: RELIABILITY**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
|---|---|
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for reliability:<br>• Student test scores have adequate levels of reliability (e.g., coefficient alpha > 0.75).<br><br>This application contains evidence of the *desired* criteria for reliability:<br>• Standard errors provided for students growth scores.<br>• Student growth classifications have adequate decision consistency.<br>• Teacher effectiveness classifications demonstrate adequate consistency.<br>*Examples include agreement statistics (e.g., kappa coefficients) based on simulation studies.* | **Check all that apply:**<br><br>☒<br><br>☒<br>☐<br>☐ |
| **2.2(D)-ii: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – ALIGNMENT**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for alignment validity:<br>• Evidence that test content is sufficiently aligned with New York State Learning Standards and covers a range of measurable standards. Documentation that demonstrates that:<br>(a) at least 80% of the test measures content aligned with NYS learning standards,<br>(b) no more than 20% of test content is aligned with other learning standards or objectives, and<br>(c) a range of content from the NYS learning standards is measured<br><br>*Note: Other relevant standards can be proposed if NYS Learning Standards do not apply to subject area.*<br><br>This application contains evidence of the *desired* criteria for alignment validity:<br>• 100% alignment between NYS Learning Standards and assessment. | **Check all that apply:**<br><br>☒<br><br>☒ |
| **2.2(D)-iii: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – RELATIONS TO OTHER VARIABLES**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for validity in relation to other variables:<br>• Evidence students' growth scores are correlated with other measures of student progress (e.g., r > .5 with measures such as the number of objectives | **Check all that apply:** |

| | |
|---|---|
| mastered by a student over the course of the year, teachers' ratings of students' progress, or scores from other assessments).<br><br>This application contains evidence of the *desired* criteria for validity in relation to other variables:<br>• Evidence teacher effectiveness ratings are positively correlated (e.g., r > .5) with other measures of teaching effectiveness. | ☒<br><br>☐ |
| **2.2(D)-iv: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – INTERNAL STRUCTURE**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for validity of internal structure:<br>• Scale properties appropriate for growth model used (*see notes*). Total scores and subscores on student assessments should be supported by dimensionality analyses (e.g., IRT residual analyses, factor analyses).<br><br>This application contains evidence of the *desired* criteria for validity of internal structure:<br>• Evidence students' scores are on an interval scale.<br><br>**Notes: If **gain score model** is used, evidence is needed that students' pretest and posttest scores are on the same scale. If **student growth percentile** model used, justification for the number of years included in the model should be provided. If **growth-to-proficiency**, projection, or **value-added models** are used, evidence is needed that the model explains a significant amount of variability in student achievement. Also, models should demonstrate robustness to missing data.* | **Check all that apply:**<br><br>☒<br><br>☒ |
| **2.2(D)-v: Technical Documentation Related to Assessment and Student Growth Score Properties: UTILITY AND COMPREHENSIBILITY**<br>***Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.*** | |
| **For supplemental assessments used in conjunction with growth models:**<br>This application contains evidence of the *minimum* criteria for utility and comprehensibility:<br>• Technical documentation that describes how student growth and educator effectiveness are calculated.<br><br>This application contains evidence of the *desired* criteria for utility and comprehensibility:<br>• Student growth reports support instructional improvement. Resources and supporting materials available to the field. | **Check all that apply:**<br><br>☒<br><br>☒ |
| **2.2(E)-i: Technical Documentation Related to Aggregating Student-Level Growth Scores to Teacher-Level Scores: CREATION OF TEACHER LEVEL SCORES** | |

| | |
|---|---|
| **For supplemental assessments used in conjunction with growth models:**<br>This application includes a narrative description of how student-level scores are aggregated to create a single teacher-level score for each teacher. | ☒ ☐ N/A |
| **2.2(E)-ii: Technical Documentation Related to Aggregating Student-Level Growth Scores to Teacher-Level Scores: EXCLUSION RULES** | |
| This application includes a description of any exclusion rules that remove students associated with a given teacher from the teacher's teacher-level score (either through a growth model or in conjunction with an SLO). | ☒ ☐ N/A |
| **2.2(F): Technical Documentation Related to Converting Teacher-Level Growth Score to New York State's 0-20 APPR Scale** | |
| This application includes a crosswalk that maps scores on the assessment's aggregated teacher-level growth score to the required New York State teacher and principal evaluation metric, which ranges from 0-20.<br><br>This application includes procedures for converting teacher-level growth scores to the 0-20 APPR scale comply with the New York Standards for each evaluation rating category, which are based on the following definitions.<br><br>**For supplemental assessments used in conjunction with growth models:**<br>This application includes an explanation of the assignment of HEDI rating categories based on the following ranges:<br>• Highly Effective: results are well-above State average* for similar students<br>• Effective: results meet State average* for similar students<br>• Developing: results are below State average* for similar students<br>• Ineffective: Results are well-below State average* for similar students | ☒<br><br>☒<br><br>☒ ☐ N/A |
| **2.2(G)-i: Technical Documentation Related to Fairness: TEST TAKERS**<br>Consistent with the new Testing Standards (2014), there is an increased focus in the industry on fairness of assessments and their uses. Please provide evidence of fairness for both the proposed assessment and, if applicable, the proposed growth model. | |
| This application includes evidence that the proposed assessments are fair to all test takers (e.g., Differential Item Functioning [DIF] / bias information, fairness evaluation / sensitivity review plan.) | ☒ |
| **2.2(G)-ii: Technical Documentation Related to Fairness: TEACHER GROWTH SCORES** | |
| This application includes evidence of fairness of the proposed aggregated teacher growth scores (e.g., lack of correlation between aggregated teacher growth scores and student demographics).<br><br>The evidence of fairness of the proposed aggregated teacher growth scores includes an explanation of how the growth model incorporates (a) prior academic history, (b) poverty, (c) students with disabilities, and (d) English language learners. | ☒<br><br>☒ ☐ N/A |

**To be completed by the Copyright Owner/Assessment Representative of the assessment being proposed and, where necessary, the co-applicant LEA:**

| | |
|---|---|
| The Northwest Evaluation Association (NWEA)<br>1. Name of Organization (PLEASE PRINT/TYPE) | 4. Signature of Authorized Representative (PLEASE USE **BLUE** INK) |
| Geri Cohen<br>2. Name of Authorized Representative (PLEASE PRINT/TYPE) | 09/30/2015<br>5. Date Signed |
| Vice President of Finance & CFO<br>3. Title of Authorized Representative (PLEASE PRINT/TYPE) | |

| | |
|---|---|
| NA<br>1. Name of LEA (PLEASE PRINT/TYPE) | 4. Signature of School Representative (PLEASE USE **BLUE** INK) |
| 2. School Representative's Name (PLEASE PRINT/TYPE) | 5. Date Signed |
| 3. Title of School Representative (PLEASE PRINT/TYPE) | |