

Georgia’s Response to the United States Department of Education’s Interim Feedback Letter

This document provides Georgia’s response to the interim feedback letter received from the U.S. Department of Education on March 8, 2019 regarding Georgia’s Innovative Assessment Demonstration Authority (IADA) application. Responses are presented under each regulatory and selection criteria requirement for which additional information was requested. Responses are provided for the Cobb County School District (CCSD), Georgia MAP Assessment Partnership (GMAP), Putnam Consortium, and State of Georgia, where applicable.

Regulatory Requirement	Required information from the SEA
<p>Innovative assessment system. A demonstration that the innovative assessment system does or will--</p> <p>(1) Meet the requirements of section 1111(b)(2)(B) of the Act, except that an innovative assessment--</p> <p>(i) Need not be the same assessment administered to all public elementary and secondary school students in the State during the demonstration authority period described in 34 CFR 200.104(b)(2) or extension period described in 34 CFR 200.108 and prior to statewide use consistent with 34 CFR 200.107, if the innovative assessment system will be administered initially to all students in participating schools within a participating LEA, provided that the statewide academic assessments under 34 CFR 200.2(a)(1) and section 1111(b)(2) of the Act are administered to all students in any non-participating LEA or any non-participating school within a participating LEA; and</p> <p>(ii) Need not be administered annually in each of grades 3-8 and at least once in grades 9-12 in the case of reading/language arts and mathematics assessments, and at least once in grades 3-5, 6-9, and 10-12 in the case of science assessments, so long as the statewide academic assessments under 34 CFR 200.2(a)(1) and section 1111(b)(2) of the Act are administered in any required grade and subject under 34 CFR 200.5(a)(1) in which the SEA does not choose to implement an innovative assessment.</p>	<p>For the Georgia MAP Assessment Partnership (GMAP) assessment model, clarification is needed regarding the relationship of growth scores described in the application and the requirement that assessments yield an annual summative determination of proficiency of the State’s academic content standards.</p>

Georgia MAP Assessment Partnership

The adaptive through-year blueprint results in the creation of two different scales that will be produced from the GMAP Through-Year assessment:

- End-of-Year Summative Proficiency scores, and associated performance levels (Beginning Learner; Developing Learner; Proficient Learner; Distinguished Learner) generated based on coverage of the Georgia Milestones blueprints.
- Interim Growth Scores based on the MAP Growth RIT Scale, which will be used to identify the instructional needs of students and inform instruction throughout the year.

The GMAP summative scores will be based on summative items that represent the Georgia Standards of Excellence to the same extent as the Georgia Milestones test blueprints. The final summative score comes from two components: a) an adaptive component containing machine-scoreable items aligned to the Georgia standards (administered in the fall, winter and spring), and b) a summative performance task component administered in the spring, which will be scored by human scorers (utilizing a professional scoring vendor who specializes in performance tasks). The results of both components ensure the Milestones blueprint has been covered and are combined to create a single summative score used for annual determinations of achievement.

The adaptive algorithm is such that it ensures coverage in terms of number of items and points for component areas proportional to the coverage of the Georgia Milestones blueprints and provides students with multiple opportunities (e.g. fall, winter, spring) to demonstrate progress towards proficiency against these blueprints as the year progresses. The summative score comes from a combination of machine-scorable items administered throughout the year and standardized writing performance tasks administered in the spring and scored by a professional hand-scoring vendor. The machine-scorable items include, but are not limited to, multiple select, gap match, and graphic gap match, drag and drop, text entry, and composite items. Composite items might have parts of more than one type such as an Evidence Based Selected Response item type in ELA. All items used in the summative score will be aligned to the state standards.

Although formative performance tasks will be available for administration throughout the school year, the summative performance tasks will only be administered at the end of the year under standardized administration conditions. The adaptive component of machine-scorable items will be administered in the fall, winter, and spring under standardized administration conditions, using a lock-down browser.

It is important to note that the items and adaptive engine used to generate the summative scores are not the same as those used for our currently available MAP® Growth™ assessment. The GMAP through-year adaptive engine is a constraint-based engine that enables the assessment to adapt to cover the testing blueprint, providing coverage of on-grade content over the course of the year. A similar version of the constraint-based engine has successfully been used in Nebraska's end-of-year summative testing system, the Nebraska Student-Centered Assessment System (NSCAS).

Interim Growth scores will be derived from a subset of items, will be reported immediately after each test event, and will align to the RIT scale to provide normative growth information. Interim tests will also provide projections to the summative scale in the fall and winter based on the subset of on-grade items aligned to the Georgia Standards of Excellence, giving teachers and students early warning signals if a student is not on track to proficiency. Summative machine-scoreable items are aligned to the Georgia Standards of Excellence, and the proportion of items to standards will match the proportions of content in the summative blueprint for Georgia Milestones. Summative Performance Task items administered at the end of the year, aligned to the Georgia Standards of Excellence ensure full coverage of the writing blueprint.

Table 1: Summary of Scores and Uses

	Fall	Winter	Spring	Notes
Scores				
Projected Scale Score	X	X	X	Intended to help educators understand how well students are accessing grade level content throughout the year
Final Scale Score			X	End-of-Year Summative Proficiency score, and associated performance levels (Beginning Learner; Developing Learner; Proficient Learner; Distinguished Learner) generated based on coverage of the Georgia Milestones blueprints (upon completion of scoring).
Growth Score Towards Proficiency	X (if prior year scores are available)	X	X	Intended to help educators understand student growth towards grade-level proficiency.
Normative Growth Score/ RIT	X	X	X	Intended to help educators understand historical student growth regardless of student proficiency level.
Item Sets and Types				
Summative Items to grade-level blueprint	X	X	X	Full adaptive blueprint covered by end of final test. Test prioritizes on-grade content and blueprint coverage. Only these items used in projected scale score and these items plus summative performance task for calculation of final scale score.
Interim Items as appropriate	X	X	X	Used in Normative Growth Scale/RIT.
Optional Interim Performance Tasks, Teacher Scored	X	X		Used by educators to elicit deeper thinking between assessments.
Summative Performance Tasks			X	Constructed response items similar to those on Georgia Milestones, used for full coverage of summative blueprint, covers writing and math standards not included in machine scored items.

Regulatory Requirement	Required information from the SEA
(2)(i) Align with the challenging State academic content standards under section 1111(b)(1) of the Act, including the depth and breadth of such standards, for the grade in which a student is enrolled; and (ii) May measure a student’s academic proficiency and growth using items above or	For all three proposed models, the Cobb County School District (CCSD); the GMAP; and the Putnam County consortium: 1. Information regarding the processes and procedures of their multiple event administration designs, in order to ensure that all students who participate in the pilot

<p>below the student’s grade level so long as, for purposes of meeting the requirements for reporting and school accountability under sections 1111(c) and 1111(h) of the Act and paragraphs (b)(3) and (b)(7)-(9) of this section, the State measures each student’s academic proficiency based on the challenging State academic standards for the grade in which the student is enrolled;</p> <p>(3) Express student results or competencies consistent with the challenging State academic achievement standards under section 1111(b)(1) of the Act and identify which students are not making sufficient progress toward, and attaining, grade-level proficiency on such standards;</p>	<p>assessments are assessed against all of the State’s academic content standards (e.g., what are the procedures in the event of a student absence from one of the scheduled testing administrations?).</p> <p>2. Information regarding the content specifications of the pilot assessments. Specifically, prototype test blueprint documents comparable to those used for the statewide assessments must be provided to ensure that the pilots are designed in a way to assess the full depth and breadth of the academic content standards.</p> <p>For the CCSD, a plan to express student results in terms of the State’s academic achievement standards (e.g., what are the psychometric linking designs/decisions inherent in the scaling plan for the CCSD test? What justifications are there to support the scaling plan for the test?)</p> <p>For the GMAP, clarification is needed regarding the relationship of growth scores described in the application and the requirement that assessments yield an annual summative determination of proficiency of the State’s academic content standards. Information provided for the GMAP in (b)(1) will also address the concern raised in this requirement.</p>
---	--

Cobb County School District

Assessments will be administered online only during specific administration windows each quarter. Students who are absent on the assessment administration date will be required to take the assessment at a later date during the regular assessment administration window or during a specified make-up assessment administration window which will be scheduled immediately following the regular assessment administration dates. A non-participation list will be generated by the assessment administration platform for each test which will be used by teachers to schedule make-up testing. This will ensure all students who are required to participate in the assessments will have the opportunity to do so.

Prototypes of assessment blueprints for all grades and content areas that will be tested in the first year of the assessment administration pilot are included in Appendix A. The prototype blueprints were developed using the CCSD Instructional Frameworks and Georgia Milestones assessment blueprints as their foundation to ensure all GTLS assessments meet the depth and breadth of the challenging state academic standards, the Georgia Standards of Excellence (GSE).

CCSD expects the results of the through-year assessments to be provided at the standards levels to students and teachers to allow for instruction to be adjusted throughout the year. However, to arrive at the required end-of-year achievement designation for each student, CCSD expects to combine the raw scores of the multiple through-year assessments into a single raw score that can be then scaled to be comparable to the State assessments. The exact details of the procedures and the plans for the comparability studies will be developed in consultation with our external psychometric partners. CCSD has initiated the process of securing external partners. The qualifications of the psychometric partners that are sought are detailed in a subsequent response.

Georgia MAP Assessment Partnership

The GMAP summative scores will be based on summative items that represent the Georgia Standards of Excellence to the same extent as the Georgia Milestones test blueprints. The final summative score comes from two components: a) an adaptive component containing machine-scoreable items aligned to the Georgia standards (administered in the fall, winter, and spring), and b) a summative performance task component administered in the spring, which will be scored by human scorers (utilizing a professional scoring vendor who specializes in performance tasks). The results of both components ensure the Milestones blueprint has been covered and are combined to create a single summative score used for annual determinations of achievement. The adaptive algorithm is such that it ensures coverage in terms of number of items and points for component areas proportional to the coverage of the Georgia Milestones blueprints and provides students with multiple opportunities (e.g. fall, winter, spring) to demonstrate progress towards proficiency against these blueprints as the year progresses. GMAP blueprints will mirror the Georgia Milestones in terms of content coverage and weights. For instance, the screen shot below shows the content domains, standards, and weights for grade 3 ELA taken from the Georgia Milestones Assessment Guide for Grade 3. The GMAP grade 3 ELA summative scores will target the percentages shown in this table with a small margin of error.

Figure 1. Georgia Milestones Grade 3 Domain Structure for ELA
GRADE 3 ENGLISH LANGUAGE ARTS (ELA): DOMAIN STRUCTURES AND CONTENT WEIGHTS

Domain	Standard		Approximate Weight
Reading and Vocabulary	ELAGSE3RI1	ELAGSE3RL3	53%
	ELAGSE3RI2	ELAGSE3RL4	
	ELAGSE3RI3	ELAGSE3RL5	
	ELAGSE3RI4	ELAGSE3RL6	
	ELAGSE3RI5	ELAGSE3RL7	
	ELAGSE3RI6	ELAGSE3RL9	
	ELAGSE3RI7	ELAGSE3L4	
	ELAGSE3RI8	(4a, 4b, 4c, 4d)	
	ELAGSE3RI9	ELAGSE3L5	
	ELAGSE3RL1	(5c)	
	ELAGSE3RL2		
Writing and Language	ELAGSE3W1	ELAGSE3L1	47%
	(1a, 1b, 1c, 1d)	(1a, 1b, 1c, 1d, 1e, 1f, 1g, 1h, 1i)	
	ELAGSE3W2	ELAGSE3L2	
	(2a, 2b, 2c, 2d)	(2a, 2b, 2c, 2d, 2e, 2f, 2g)	
	ELAGSE3W3	ELAGSEL3	
	(3a, 3b, 3c, 3d)	(3a)	
	ELAGSE3W4		
	ELAGSE3W7		
ELAGSE3W8			

The summative score comes from a combination of machine-scorable items administered throughout the year and standardized writing performance tasks, administered in the spring and scored by a professional hand-scoring vendor. The machine scoreable items include multiple choice and technology-enhanced items including, but not limited to multiple select, gap match, and graphic gap match, drag and drop, text entry, and composite items which might have parts of more than one type such as an Evidence Based Selected Response item type in ELA all of which are aligned to the state standards.

It is important to note that the items and adaptive engine used to generate the summative scores are not the same as those used for our currently available MAP Growth assessment. The GMAP through-year adaptive engine is a constraint-based engine that will enable the assessment to adapt to cover the testing blueprint, providing coverage of on-grade content. Once the blueprint requirements have been satisfied, off-grade items will be administered to students if the adaptive algorithm determines that is the next best item for a student. A similar version of the constraint-based engine has successfully been used in Nebraska's end-of-year summative testing system the NSCAS.

The adaptive component is controlled by a constraint engine that selects items to fully and proportionally represent the blueprint by the end of the final test administration. Each test event – fall, winter, and spring – will be managed by the constraint engine to match the appropriate test blueprint for each grade level assessed. To be clear, this model is NOT a “through-course summative model” that dictates pacing and divides the content standards into separate blueprints and distributes them across the interim tests. Instead, the through-year model uses a constraint-based adaptive engine that provides content representation across administrations such that the full blueprint is assessed by the conclusion of the spring test event. The breadth of the content standards in the blueprint are repeatedly measured across the interim tests, but the depth (DOK, cognitive rigor, complexity, and context of the items) will vary according to the individual student's ability level. If a student has not experienced the full breadth and depth in the fall and winter assessment (or missed one or more of the earlier assessments and has been unable to take a make-up, or has come in mid-year), the constraint-based engine recognizes this gap and the final spring test event will present the items necessary to satisfy the blueprint. Once the blueprint requirements have been satisfied, off-grade items will be administered to students if the adaptive algorithm determines that is the next best item for a student. For example, if a student is very advanced for their grade, they may demonstrate mastery of content standards at the Distinguished Learner level prior to spring test administration. In such a case, the spring test event will adapt to the most challenging on-grade items and even adapt above grade if needed. If a student is behind their grade level, the test may adapt to the student's instructional level, even below grade, in the fall and winter and they may not be exposed to all items from the full blueprint until the final spring test event. In this situation, the spring test event for the student will be less adaptive because the constraint engine will ensure that the full grade-level blueprint and content domain is measured during that event. All student scores will not be official and final until the writing performance task items have been scored by the scoring vendor and combined with the adaptive component of the test.

Similar to the interim tests currently being given in GMAP districts, tests in the fall, winter, and spring will be scheduled at the local level, within a designated fall, winter, and spring testing window, with the ability to give students a make-up opportunity within the test window as needed. If a student demonstrates a low, medium, or high mastery over the content in a previous administration, the system will pick up where the student left off and continue the assessment of grade-level content. If a student is absent in the fall or winter test event or both and did not make up those test events, the adaptive engine will assess the student on the full blueprint before the final spring test event is finished. In this

case, the last test event may be slightly longer and slightly less adaptive if necessary to provide blueprint coverage.

Putnam Consortium

The Navvy assessment system is an on-demand, or as-needed, assessment system. As such, no testing windows are pre-determined by the Putnam Consortium or by Navvy Education. Teachers of a class section may open a testing window to administer an assessment as needed during times when the teacher is available to proctor the assessment. School- and/or district-administrators may also open a testing window to administer an assessment. These testing windows are determined by the district, though the district may pass this level of control down to the school- or teacher-level. Thus, assessment remains fluid so that it can occur when information is needed regarding the student competency of a standard. As a specific example, if a student is absent and misses an assessment, then the teacher may opt to administer the assessment at a time that is most suitable for that individual student. This suitable time may be the next day they return to school: perhaps during class where differentiated learning and instruction is taking place, during special periods that are purposed for flexible instruction or remediation, or before/after school. Or the teacher may determine that another day is more suitable for the student given the context.

Please see Proprietary Appendix B for content specifications and prototype blueprint documents.

Regulatory Requirement	Required information from the SEA
<p>(4)(i) Generate results, including annual summative determinations as defined in paragraph (b)(7) of this section, that are valid, reliable, and comparable for all students and for each subgroup of students describe in 34 CFR 200.2(b)(11)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act, to the results generated by the State academic assessments described in 34 CFR 200.2(1)(1) and section 1111(b)(2) of the Act for such students.</p> <p>Consistent with the SEA’s of consortium’s evaluation plan under 34 CFR 200.106(3), the SEA must plan to annually determine comparability during each year of its demonstration authority period in one of the following ways:</p> <p>(A) Administering full assessments from both the innovative and statewide assessment systems to all students enrolled in participating schools, such that at least once in any grade span (i.e., 3-5, 6-8, or 9-12) and subject for which there is an innovative assessment, a statewide assessment in the same subject would also be administered to all such students. As part of this determination, the innovative assessment and statewide</p>	<p>For the CCSD:</p> <ol style="list-style-type: none"> 1. Information regarding the sample sizes planned in the concurrent testing of both the pilot and the State summative assessments. 2. An explanation for how variations across forms (that may differ in both mode, content and item types) will: <ol style="list-style-type: none"> a. Be included in the comparability analyses for the CCSD pilot assessments. b. Will result in valid and reliable estimates of student achievement for all students that participate in the pilot assessments. 3. An assurance that CCSD will perform comparability analyses for each year that pilot assessment scores are used in lieu of State assessment scores during the period of the authority (if awarded). <p>For the GMAP:</p> <ol style="list-style-type: none"> 1. An assurance that GMAP will perform comparability analyses for every year that pilot assessment scores are used in lieu of State assessment scores during the period of the authority (if awarded). 2. More detailed and specific information

assessment need not be administered to an individual student in the same school year.

(B) Administering full assessments from both the innovative and statewide assessment systems to a demographically representative sample of all students and subgroups of students described in section 1111(c)(2) of the Act, from among those students enrolled in participating schools, such that at least once in any grade span (i.e., 3-5, 6-8, or 9-12) and subject for which there is an innovative assessment, a statewide assessment in the same subject would also be administered in the same school year to all students included in the sample.

(C) Including, as a significant portion of the innovative assessment system in each required grade and subject in which both an innovative and statewide assessment are administered, items or performance tasks from the statewide assessment system that, at a minimum, have been previously pilot tested or field tested for use in the statewide assessment system.

(D) Including, as a significant portion of the statewide assessment system in each required grade and subject in which both an innovative and statewide assessment are administered, items or performance tasks from the innovative assessment system that, at a minimum, have been previously pilot tested or field tested for use in the innovative assessment system.

(E) An alternative method for demonstrating comparability that an SEA can demonstrate will provide for an equally rigorous and statistically valid comparison between student performance on the innovative assessment and the statewide assessment, including for each subgroup of students described in 34 CFR 200.2(b)(11)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act;

(ii) Generate results, including annual summative determinations as defined in paragraph (b)(7) of this section, that are valid, reliable, and comparable, for all students and for each subgroup of students described in 34 CFR 200.2(b)(11)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act, among participating schools and LEAS in the innovative assessment demonstration authority.

regarding the inclusion of performance tasks on the GMAP assessment. Information provided in response to (b)(2) above may address this request.

3. A description of how variations across forms of the GMAP (e.g., forms with performance tasks and forms without performance tasks) will be analyzed for comparability.
4. A description of how the results of various linking studies described in the application will be estimate comparability.
5. A clear indication of how comparability will be determined (i.e., one of the five methods identified in the regulations and shown on the left).

Consistent with the SEA’s or consortium’s evaluation plan under 34 CFR 200.106(e), the SEA must plan to annually determine comparability during each year of its demonstration authority period;	
--	--

Cobb County School District

The planned sample sizes and overall comparability studies will be determined in consultation with CCSD’s external psychometric partners. In the first year of the demonstration authority, CCSD expects that all district students will participate in concurrent testing. In subsequent years, CCSD expects that students participating in concurrent testing will gradually decrease depending on psychometric requirements. Likewise, details of the comparability, reliability, and validity studies of the CCSD pilot assessments are in the developing stages in consultation with the external psychometric partners.

CCSD expects the pilot assessments to be delivered in a single mode (online) and use the same content and item types to all students within a grade and subject.

CCSD will perform comparability analyses every year of the period of the authority. A goal of CCSD will be to minimize the burden of concurrent testing on our schools within the limits of demonstrating the comparability between the pilot assessments and the existing State assessments. The comparability plans will be developed in consultation with our external psychometric partners and any guidance provided by the Georgia Department of Education.

Georgia MAP Assessment Partnership

Assurance that GMAP will perform comparability analyses each year

In addition to the comparability study used in the initial pilot year where students will take both the Georgia Milestones and the GMAP assessment, additional comparability studies will be conducted in subsequent years. The GMAP consortium will perform additional comparability studies every year (Years 4 and 5) that the pilot assessment scores are in use in lieu of the State assessment scores.

Performance Tasks and comparability

Summative Performance tasks to ensure coverage of the blueprint will be given in the spring and scored by professional scoring vendors. There will not be multiple forms of the assessment with and without these tasks, so no form comparability studies are needed as there are not multiple forms. Both the summative items and summative performance tasks given in the spring will be used to ensure blueprint coverage for creation of the summative score. Scores from fall and winter interim tests will not be used as final summative decisions; however, in the spring test event, the initial ability estimate that the adaptive algorithm will use to select items will come from the ability estimate from the last interim test event. The test will then adapt based on the student’s responses. Like other adaptive tests, the GMAP adaptive test will maintain comparability by complying with the summative blueprint, test specifications, and test administration policies.

Description of linking and comparability studies

Year 3: Option A will be used. In Year 3 for English language arts (ELA) and Mathematics and Year 4 for Science, we plan to conduct comparability studies by implementing Option A of section 4(i) "Administering full assessments from both the innovative and statewide assessment systems to all students enrolled in participating schools, such that at least once in any grade span (i.e., 3-5, 6-8, or 9-12) and subject for which there is an innovative assessment, a statewide assessment in the same subject would also be administered to all such students." We will use equipercentile linking to create concordance tables and will set performance standards (cut scores) that minimize achievement level misclassification across the two scales. These cut scores will be used until standard setting in Year 5 establishes new performance standards. The performance standards that are established in Year 5 will be informed by the previous cut scores.

The cut scores for achievement levels will be established in Year 3 for ELA and Mathematics, and Year 4 for Science using Option A with the equipercentile linking method.

Years 4 and 5: Priority – Option C In years 4 and 5, the initial plan is to use option C from the regulations to ensure comparability is maintained. If the GaDOE does not approve the use of Georgia Milestones items for option C, Option B from the regulations will be used.

In Years 4 and 5 for ELA and Mathematics, and Year 5 in Science, the priority is to apply Option C in section 4(i) "Including, as a significant portion of the innovative assessment system in each required grade and subject in which both an innovative and statewide assessment are administered, items or performance tasks from the statewide assessment system that, at a minimum, have been previously pilot tested or field tested for use in the statewide assessment system." We interpret this design to be consistent with a "common item non-equivalent groups design", in which different tests measuring the same construct can be linked together¹. The consortium is discussing with the Georgia Department of Education (GaDOE) the potential to use previously field-tested items, representative of the Georgia Milestones blueprints, and Item Response Theory (IRT) calibrations for said items. If that is feasible, we will use IRT linking procedures to link the Georgia Milestones scale to the GMAP through-year summative assessment scale. This will allow us to produce two scale scores, one for the Georgia Milestones and one for the GMAP through-year summative. Students will be classified into achievement levels according to both scales and the classification agreement will be examined to determine the degree of comparability between the scales. Option C is our first choice for years 4 and 5 because it does not require double testing; however, because Option C requires the GaDOE to approve the use of Milestones items, there is a risk that this approval may not be granted. Therefore, option B from the regulations will be the next preferred method for comparability if GaDOE does not grant approval.

Year 4 & 5: First Choice Backup Option, Option B: An alternative to Option C, is Option B, which requires a sample of GMAP students to double test. We will use the same equipercentile linking methodology as described under option A to verify comparability. The GMAP districts have agreed to this alternative if option C is not approved by the GaDOE.

Context for Design

Item development will continue after Year three, however, the blueprints and item specifications will remain consistent across time. The scale will be established in Year 3 with items measuring the full breadth and depth of the content standards. Because the blueprints will be well represented by the

¹Kolen, M.J. & Brennan, R.L. (2004). Test Equating, Scaling, and Linking: *Methods and Practice*. New York: Springer.

items field tested in Year 3, the scale that emerges from that year will represent the full domain in each subject and grade. Thereafter, newly written items that conform to the item specifications will be embedded as field test items within the operational assessment, thereby permitting us to place the newly field-tested items onto the same scale that was established in Year 3. The blueprints and item specifications will preserve the operational definition of the constructs, while the embedded field testing of new items will ensure that items are replaced and refreshed to prevent item over-exposure.

Please note: There will be two types of performance tasks in the through-year assessment system: formative and summative. When the proposal mentioned "this is likely", that clause was referencing the formative performance tasks. To clarify, the formative assessments are "likely" to be developed, while the summative performance tasks are *certainly* going to be developed. The summative items will be scored by professional scoring services, while the formative items that are not included in the proficiency scores are intended to be scored by teachers as part of a balanced assessment system.

Regulatory Requirement	Required information from the SEA
(6) For purposes of the State accountability system consistent with section 1111(c)(4)(E) of the Act, annually measure in each participating school progress on the Academic Achievement indicator under section 1111(c)(4)(B) of the Act of at least 95 percent of all students, and 95 percent of students in each subgroup of students described in section 1111(c)(2) of the Act, who are required to take such assessments consistent with paragraph (b)(1)(ii) of this section;	For all three proposed models -- the CCSD; the GMAP; and the Putnam County consortium, more information regarding the processes and procedures of these multiple event administration designs, in order to ensure that at least 95 percent of all students are able to participate in all parts of the pilot assessments.

Cobb County School District

See response to required information requested under requirement 2(i)-(3).

Georgia MAP Assessment Partnership

GMAP districts will ensure, as they do currently for the Georgia Milestones Assessment, that at least 95 percent of all students participate in the assessment.

The adaptive component is controlled by a constraint engine that selects items to ensure that the blueprint is fully and proportionally represented by the end of the final test administration. Each test event – fall, winter, and spring – will be managed by the constraint engine to match the appropriate test blueprint for each grade level assessed. To be clear, this model is NOT a “through-course summative model” that dictates pacing and divides the content standards into separate blueprints and distributes them across the interim tests. Instead, the through-year model uses a constraint-based adaptive engine that ensures blueprint representation across administrations. The breadth of the content standards in the blueprint are repeatedly measured across the interim tests, but the depth (cognitive rigor, content, and context of the items that are associated with item difficulty) will vary according to the individual student’s ability level. If a student has not experienced the full on-grade depth in the fall and winter assessment (or missed one or more of the earlier assessments and been unable to make-up the assessment, or has come mid-year), the constraint-based engine recognizes this gap and the final spring

test event will present the items necessary to satisfy Federal summative guidelines. As students’ progress through stages of more advanced thinking in the content, the system picks up where the student left off in the previous administration and continues the assessment of grade-level content in the next administration. Once the one-grade blueprint requirements have been satisfied, off-grade items will be administered to students if the adaptive algorithm determines that is the next best item for a student. For example, if a student is very advanced for their grade, they may demonstrate mastery of content standards at the Distinguished Learner level prior to spring test administration. In such a case, the spring test event will adapt to the most challenging on-grade items and even adapt above grade if needed. If a student is behind their grade level, the test may adapt to the student’s ability level, even below grade, in the fall and winter to inform instructional next steps. These students may not be exposed to all items from the on-grade until the final spring test event. In this situation, the constraint engine will ensure that the full blueprint and content domain is measured with grade level items during that event to produce the proficiency score.

Similar to the interim tests currently being given in GMAP districts, tests in the fall, winter, and spring will be scheduled at the local level, within a designated fall, winter, and spring testing window, with the ability to give students a make-up opportunity during the testing window as needed. If a student demonstrates a low, medium, or high mastery over the content in a previous administration, the system will pick up where the student left off and continue the assessment of grade-level content. If a student is absent in the fall or winter test event or both and did not make up those test events, the adaptive engine will assess the student on the full blueprint before the final spring test event is finished. In this case, the last test event may be slightly longer and slightly less adaptive, if necessary, to ensure blueprint coverage. This will allow all students to be assessed against the entirety of the state blueprint for their grade.

Putnam Consortium

As with Georgia’s current state-level testing, at least 95% of students in participating districts will participate in the innovative pilot assessments.

To help participating districts monitor participation, Navvy will provide a dashboard that tracks student participation data at the class-, school-, and district-levels in real time. Navvy will also provide suggested implementation benchmarks for through-year participation that districts may use as a guide for monitoring implementation.

For affiliate districts, student participation data will be used as an indicator of a district’s preparedness for transitioning to a participating district with respect to the district’s promise to be successful in having at least 95% of their students participate in the pilot assessments.

Regulatory Requirement	Required information from the SEA
(7) Generate an annual summative determination of achievement, using the annual data from the innovative assessment, for each student in a participating school in the demonstration authority that describes-- (i) The student’s mastery of the challenging State academic standards under section 1111(b)(1) of	For all three proposed models -- the CCSD; the GMAP; and the Putnam County consortium, as noted in (b)(2) above, information is needed regarding the processes and procedures of these multiple event administration designs, in order to ensure that all students who participate in pilot assessments are assessed against all of the

<p>the Act for the grade in which the student is enrolled; or (ii) In the case of a student with the most significant cognitive disabilities assessed with an alternate assessment aligned with alternate academic achievement standards under section 1111(b)(1)(E) of the Act, the student’s mastery of those standards;</p>	<p>academic content standards in determining the annual summative determination. For the CCSD and GMAP models, clarification regarding as to how the “through-year” assessments will be aggregated into one summative determination.</p>
--	---

Cobb County School District

See response to required information requested under requirement 2(i)-(3).

Georgia MAP Assessment Partnership

The GMAP summative scores will be based on summative items that represent the Georgia Standards of Excellence to the same extent as the Georgia Milestones test blueprints. The final summative score comes from two components: a) an adaptive component containing machine-scoreable items aligned to the Georgia standards, and b) a summative performance task component administered in the spring, which will be scored by human scorers (utilizing a professional scoring vendor who specializes in performance tasks). The results of both components ensure the Milestones blueprint has been covered and are combined to create a single summative score used for annual determinations of achievement. To ensure comparability to the Milestones blueprints, the number of items and points in each component will be proportional to the Milestones blueprint. This helps support the comparability of the scores from the two systems. A student’s record of item responses will be collected and stored in memory so that when the student returns to take the next test (winter or spring), the adaptive engine will know what content the student had previously been tested on. The adaptive engine will then continue where the student left off, adapting the test to determine the student’s proficiency relative to grade-level standards. Students will be given repeated opportunities (e.g. fall, winter, spring administrations) to demonstrate proficiency on the adaptive component in the winter and/or spring administrations.

The adaptive component is controlled by a constraint engine that selects items to fully and proportionately represent the blueprint by the final test administration. Each test event – fall, winter, and spring – will be managed by the constraint engine to match the appropriate test blueprint for each grade level assessed. The system will continue where the student left off, but if a student is absent on the fall or winter test event or both and is unable to make up the test during the assessment window, the system will recognize this and will ensure that the blueprint is fully covered by the spring administration. Once the blueprint requirements have been satisfied, off-grade items will be administered to students if the adaptive algorithm determines the student is off-grade. It is important to note that the items and adaptive engine used to create the summative score are not the same as those used in our currently available MAP Growth assessment. The GMAP through-year adaptive engine is a constraint-based engine that will provide coverage of on-grade content according to the Georgia Milestones blueprints. This constraint-based engine has successfully been used in Nebraska’s end-of-year summative testing system, NSCAS.

Putnam Consortium

Please see Proprietary Appendix B for content specifications and prototype blueprint documents.

Regulatory Requirement	Required information from the SEA
<p>(8) Provide disaggregated results by each subgroup of students described in 34 CFR 200.2(b)(11)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act, including timely data for teachers, principals and other school leaders, students, and parents consistent with 34 CFR 200.8 and section 1111(b)(2)(B)(x) and (xii) and section 1111(h) of the Act, and provide results to parents in a manner consistent with paragraph (b)(4)(i) of this section and part 200.2(3);</p>	<p>For the GMAP model:</p> <ol style="list-style-type: none"> 1. A description of how the time needed for anticipated analyses (such as linking activities) is accounted for between operational testing and reporting. Without a clear understanding of these procedures, it is unclear that the pilot assessment results will be able to be produced disaggregated for all required sub-groups. 2. Clarification is needed regarding the relationship of growth scores described in the application and the requirement that assessments yield an annual summative determination of proficiency of the State’s academic content standards (as noted in b(1) above). <p>For the CCSD model, as noted under (b)(2) above, information describing a plan (which would address psychometric linking designs/decisions and provide justifications) to express student results in terms of the State’s academic achievement standards is needed in order to report results for all required sub-groups.</p>

Cobb County School District

See response to required information requested under requirement 2(i)-(3).

Georgia MAP Assessment Partnership

Description of time needed between operational testing and reporting

In Year 3 of the pilot, participants will administer both the GMAP adaptive, through-year assessment and the Georgia Milestones assessment. The Georgia Milestones assessment will be used for accountability purposes and a subsequent study will be conducted to calibrate the GMAP item pool and pre-scale GMAP results for Year 4 administration. As an adaptive test, all the operationally scored items are pre-scaled and calibrated, allowing for scale scores to be produced immediately once the items are scored. No linking or equating procedures need to be performed on this portion of the test; however, performance tasks will need to be scored and combined with the adaptive portion of the test in order to produce a final summative score. NWEA will be using a professional scoring vendor to hand score all the performance tasks. These performance tasks will be collected within our digital system, and the vendor will produce scores and return them to NWEA quickly enough to permit psychometricians to combine

the item scores into the appropriate files for scoring. This will be done in a timeframe similar to reporting time constraints for Georgia Milestones, with preliminary results of the hand-scored items anticipated in a 2-3 week window, and final summative determinations returned within a 6-8 week timeline. Results of the assessment will be provided to Georgia districts at the student level, permitting data transfers to the state that allow for inclusion in the state accountability system for all students, student groups, and subgroups. GaDOE will best determine how growth scores will be handled, however the conversion table that comes from the year 3 linking study can be used to find a Milestones scale score for any GMAP summative score. This Milestone’s ‘equivalent’ score could be used by GaDOE when producing the student growth percentiles (SGP) used in the CCRPI.

In Year 3, an equipercentile linking study will be done between the GMAP summative scale and the Georgia Milestones using a common person design. This will be done to establish cut scores on the GMAP summative scale that closely correspond to the Georgia Milestones test and produce maximally consistent classification decisions. These cut scores will be carried forward on the GMAP summative scale in Years 4 and 5. Comparability will be established at the end of year 3, and to meet the requirements of the pilot for USED, additional comparability studies will be conducted using options that do not require all students double-test. Comparability studies will be done in the summers of Years 3, 4, and 5; however, in Years 4 and 5, comparability studies and linking studies will *not* have to be performed before scores are reported because the cut scores will be established in Year 3 using the equipercentile linking results as previously mentioned.

Also, it should be clarified that while the final summative score will be produced in 6-8 weeks of the end of the state test window, NWEA will provide interim scores immediately following each test administration. This information will be used by teachers for instructional feedback but is not to be confused with the summative test score, even if highly correlated.

Clarification on growth scores versus summative scores

Please see the answer to b(l) which answers this question.

Regulatory Requirement	Required information from the SEA
<p>(9) Provide an unbiased, rational, and consistent determination of progress toward the State’s long-term goals for academic achievement under section 1111(c)(4)(A) of the Act for all students and each subgroup of students described in section 1111(c)(2) of the Act and a comparable measure of student performance on the Academic Achievement indicator under section 1111(c)(4)(B) of the Act for participating schools relative to non-participating schools so that the SEA may validly and reliably aggregate data from the system for purposes of meeting requirements for--</p> <p>(i) Accountability under sections 1003 and 1111(c) and (d) of the Act, including how the SEA will identify participating and non-participating</p>	<p>For the CCSD and GMAP models, a description regarding how the GaDOE will identify participating schools in a consistent manner for comprehensive and targeted support and improvement (e.g., how will the State use the results from CCSD and GMAP pilot assessments within the statewide system of accountability).</p>

schools in a consistent manner for comprehensive and targeted support and improvement under section 1111(c)(4)(D) of the Act; and (ii) Reporting on State and LEA report cards under section 1111(h) of the Act.	
---	--

State of Georgia

Georgia’s school accountability system is the College and Career Ready Performance Index (CCRPI). It is comprised of five components – Content Mastery, Closing Gaps, Progress, Readiness, and Graduation Rate. Scores from the statewide assessment system are utilized in eight indicators, including four achievement indicators (ELA, mathematics, science, social studies) in Content Mastery, the Closing Gaps component (a single indicator), two growth indicators in Progress (ELA, mathematics), and the literacy (Lexile) indicator in Readiness.

The Content Mastery component includes four achievement scores (ELA, mathematics, science, and social studies) based on student performance on Georgia Milestones. The achievement scores utilize weights based on achievement level, where Beginning Learners earn 0 points, Developing Learners earn 0.5 points, Proficient Learners earn 1.0 point, and Distinguished Learners earn 1.5 points. The content areas for all three grade bands (elementary, middle, high) are weighted according to the number of state tests administered within each grade band. To satisfy the requirement that state accountability systems account for insufficient participation rates, if the participation rate for all students or a subgroup of students falls below 95%, the achievement score for that group of students is multiplied by the actual participation rate divided by 95%. This ensures the adjustment is proportional to the extent to which the 95% participation rate was not attained. The adjusted achievement scores are utilized in CCRPI calculations.

Each of the three proposed innovative assessment systems – CCSD, GMAP, and Putnam – will produce a summative score and an achievement level aligned with the four Georgia Milestones achievement levels (Beginning Learning, Developing Learner, Proficient Learner, and Distinguished Learner). Therefore, the achievement level produced by the three innovative assessment systems will be used in lieu of Georgia Milestones achievement levels for Content Mastery calculations.

All participating districts are expected to administer Georgia Milestones to all students (and those results will be used in the accountability system) until the district/consortia has demonstrated adequate comparability between its innovative assessment system and Georgia Milestones. The innovative assessment technical advisory committee described in Georgia’s full IADA application will review each district/consortia’s technical evidence related to comparability. Once adequate comparability has been established and approved by the GaDOE, the GaDOE will work with each district participating in that assessment consortia to amend their flexibility contracts with the State Board of Education (SBOE). Once the SBOE approves those amendments, those districts will be able to administer the innovative assessment system in lieu of Georgia Milestones and utilize results from the innovative assessment system in the state accountability system. Districts will be expected to continue to administer Georgia Milestones to the extent necessary to continue performing annual comparability analyses as outlined in the respective sections of this application. For any grade, content area, or course assessed by Georgia Milestones for which an innovative assessment system does not have a corresponding assessment (for

example, social studies), students will be expected to participate in Georgia Milestones and those scores will be used in accountability calculations.

The Closing Gaps component of CCRPI measures the extent to which all students and all subgroups of students are meeting their annual 3% improvement targets. For each available improvement target, 0 points are earned when performance does not improve (red flag), 0.5 points are earned when progress is made but the target is not met (yellow flag), and 1 point is earned when the target is met (green flag). Economically Disadvantaged (ED), English Learner (EL), and Students with Disabilities (SWD) subgroups can earn 1.5 points when a 6% improvement target is met. Closing Gaps calculations utilize the achievement scores calculated in the Content Mastery component for all students and each subgroup of students. Therefore, the achievement level produced by the three innovative assessment systems will be used in lieu of Georgia Milestones achievement levels for Closing Gaps calculations.

The Progress component of CCRPI includes student growth indicators in English language arts and mathematics and a measure of progress toward English language proficiency for English Learners. Georgia utilizes the Student Growth Percentile (SGP) model to calculate student growth in ELA and mathematics. SGPs describe the amount of growth a student has demonstrated relative to academically-similar students from across the state. The ELA and mathematics growth scores utilize weights based on growth level. SGPs of 1-29 earn 0 points, 30-40 earn 0.5 points, 41-65 earn 1 point, and 66-99 earn 1.5 points.

The CCSD assessment model proposed to provide a growth measure that can be used for the Progress component of CCRPI. The GMAP assessment model proposes that the Georgia Milestones “equivalent” scores from GMAP can be used as if they were from Georgia Milestones in the state’s SGP calculations. Alternatively, GMAP could generate student growth percentiles with GMAP scores using all participating districts. The Putnam assessment model plans to report the transitions between the four achievement levels as a discrete measure of growth that syncs with a standards-competency accountability system. The innovative assessment technical advisory committee (or Georgia’s Accountability Technical Advisory Committee, as determined by the GaDOE) will review the growth data produced by each innovative assessment system and its relationship to the growth data produced by Georgia Milestones to determine the extent to which the growth data is comparable. The TAC will also be asked to provide recommendations for next steps based on the outcome of those comparisons. The progress toward English language proficiency indicator is not impacted by the implementation of the three innovative assessment systems as all districts will continue to administer the ACCESS for ELLs, which is utilized for this indicator.

The Readiness component of CCRPI includes three indicators for elementary and middle schools (literacy, student attendance, beyond the core) and five indicators for high schools (literacy, student attendance, accelerated enrollment, pathway completion, college and career readiness). Only the literacy indicator utilizes data from the statewide assessment system, Georgia Milestones. The remaining indicators are not impacted by the implementation of the three innovative assessment systems. For the literacy indicator, the Georgia Milestones ELA assessment produces a Lexile score. The literacy indicator calculates the percent of students demonstrating reading comprehension at or above the midpoint of the College & Career Ready “Stretch” Lexile Band for each grade level.

The CCSD assessment model proposes to use the Reading Inventory to provide a Lexile measure that can be used in the literacy indicator of CCRPI. The GMAP assessment model will correlate the RIT scale for reading to the Lexile scale. The Putnam assessment model will not produce a Lexile score; therefore, this

indicator will not be included in CCRPI calculations for districts participating in this assessment system. The Lexile measures produced by the CCSD and GMAP assessment models will be used in lieu of the Lexile measures produced by the Georgia Milestones assessment in the literacy indicator. Districts participating in the Putnam assessment model will not have the literacy indicator.

When an indicator is not available (e.g., literacy), that indicator will not be populated in the CCRPI reports. The weight associated with that indicator will be redistributed proportionally to other indicators within the component. If a component score is not available, the weight associated with that component will be redistributed proportionally to other components. Therefore, an overall CCRPI score will still be generated even if an indicator or component is not available.

Table 2 provides a summary of how the innovative assessment scores will be utilized in place of Georgia Milestones in the state accountability system.

Overall CCRPI scores and component scores are utilized in the state's Comprehensive Support and Improvement (CSI) and Targeted Support and Improvement (TSI) calculations. CSI calculations will continue to utilize the overall CCRPI score (as described above) to identify the lowest performing 5% of schools. The identification of CSI schools based on graduation rates less than or equal to 67% are not impacted by the implementation of the three innovative assessment systems. TSI calculations will continue to utilize the CCRPI component scores (as described above) to identify schools with one or more low performing subgroups.

Table 2: Summary of CCRPI Indicators and Components and Use of Innovative Assessment Results

CCRPI Components and Indicators	Georgia Milestones	CCSD	GMAP	Putnam
Content Mastery				
ELA achievement	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level
Math achievement	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level
Science achievement	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level
Social studies Achievement	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level	Achievement weighted by achievement level
Progress				
ELA growth	Student growth percentile (GM)	CCSD growth measure	Student growth percentile (GMAP)	Achievement level transition
Math growth	Student growth percentile (GM)	CCSD growth measure	Student growth percentile (GMAP)	Achievement level transition
Progress toward English language proficiency	Not impacted	Not impacted	Not impacted	Not impacted
Closing Gaps	Based on Content Mastery scores	Based on Content Mastery scores	Based on Content Mastery scores	Based on Content Mastery scores
Readiness				
Literacy (Lexiles)	Lexile score provided as part of ELA assessment	Reading Inventory Lexile	Correlate RIT scale to Lexile scale	Measure not produced
Student attendance	Not impacted	Not impacted	Not impacted	Not impacted
Beyond the core	Not impacted	Not impacted	Not impacted	Not impacted
Accelerated enrollment	Not impacted	Not impacted	Not impacted	Not impacted
Pathway completion	Not impacted	Not impacted	Not impacted	Not impacted
College and career Readiness	Not impacted	Not impacted	Not impacted	Not impacted
Graduation Rate				
4-year	Not impacted	Not impacted	Not impacted	Not impacted
5-year	Not impacted	Not impacted	Not impacted	Not impacted

Cobb County School District

As described in the response to requirement 2(i)-(3), CCSD expects that the pilot assessments will be scaled to a final scaled-score achievement determination for each student that is comparable to the existing State assessments. This scaled-score determination should be able to be used in GaDOE’s existing accountability framework (CCRPI) that is currently used to identify schools for comprehensive and targeted support and improvement.

Georgia MAP Assessment Partnership

Results of the assessment will be provided to Georgia districts at the student level, permitting data transfers to the state that allow for inclusion in the state accountability system for all students, student groups, and subgroups. Scale scores, proficiency levels aligned to Georgia Milestones reporting categories of Beginning Learner, Developing Learner, Proficient Learner, and Distinguished Learner will be provided, allowing the state to calculate growth scores as they do for Georgia Milestones assessments. Additionally, normative growth data and Lexile® information will be returned, allowing the state to make determinations for accountability along with all the other data used in the accountability system.

This allows Georgia to utilize the GMAP through-year summative assessment results with the College and Career Ready Performance Index (CCRPI) to set school improvement targets based on the percent of students who are proficient and who are working towards proficiency for GMAP schools just as the State does for all Georgia districts and schools currently using Georgia Milestones results.

Application Selection Criteria	Required information from the SEA
(a)(2) The plan the SEA or consortium, in consultation with any external partners, if applicable, has to-- (i) Develop and use standardized and calibrated tools, rubrics, methods, or other strategies for scoring innovative assessments throughout the demonstration authority period, consistent with relevant nationally recognized professional and technical standards, to ensure inter-rater reliability and comparability of innovative assessment results consistent with 34 CFR part 200.105(b)(4)(ii), which may include evidence of inter-rater reliability; and (ii) Train evaluators to use such strategies, if applicable;	For the CCSD and GMAP models, more information regarding the development and use of standardized and calibrated tools, rubrics, methods, or other strategies for scoring innovative assessments throughout the demonstration authority period.

Cobb County School District

Professional Learning: All teachers who will be required to administer the assessments will be required to participate in training led by district Assessment and Instructional Technology staff to ensure they are able to appropriately use the assessment platform to administer the assessments, monitor student progress during assessment administration, and generate appropriate data reports. Specifically, training

will consist of familiarizing teachers with the administration process, including security requirements, (which will mirror the current Georgia Milestones administration for teachers), practice the online administration steps by completing a mock administration using the assessment platform, and practice generating data reports using mock assessment results. Consultants will be utilized to complete blind scoring of constructed response items – training will be provided by district content area experts and consist of acquainting scorers with the constructed response items, rubric(s), and scoring expectations. Qualifying exercises that require scorers to score sets of exemplars (i.e. sample/model student responses that have been scored and annotated by the trainers), with extensive discussion in between each set, will ensure interrater reliability. Scorers must score exemplars with 90% of their scores being the same or adjacent as the exemplars to qualify as a scorer. Periodic inclusion of exemplars among actual student responses will ensure continuous interrater reliability throughout the actual scoring sessions.

Georgia MAP Assessment Partnership

All items created for use in the summative portion of the assessment will undergo authentic development and review processes. All scoring will be done either through machine-scoring or by a hand-scoring vendor who will use standardized and calibrated tools, rubrics, methods and other strategies consistent with relevant nationally recognized professional and technical standards, to ensure inter-rater reliability and comparability of results, using trained raters.

Performance Tasks are an important part of the through-year solution to elicit authentic work from students. This will be a key component of the solution to assessing writing from the ELA construct. NWEA is committed to producing performance tasks and has staff in both content development and psychometrics with experience in the area. However, the consortia will also be engaging with in-state and national vendor-partners with additional expertise to help build a knowledge base for GMAP teachers and to provide development efforts that are well aligned with Georgia Standards for Excellence expectations.

In 2019, GMAP is planning a joint workshop with external vendor-partners who specialize in hand scoring and performance task development and educators in the state. The objectives of this workshop are to:

- Engage with educator stakeholders in Georgia to begin understanding how performance tasks are currently being used in their classrooms and how they may expand such usage with GMAP in a formative manner, and how this can inform the summative development efforts.
- In collaboration with Georgia educators, build the rubrics necessary to start development of performance tasks with content development staff who have expertise in this area.
- Increase capacity to produce performance tasks both within NWEA and the GMAP consortium.

Prior to the workshop, NWEA and external vendor-partner content development and psychometric staff will meet and collaborate to finalize an agenda and process, including roles and responsibilities and precise outcomes, to be used for and gained from the meeting. Prework will include the development of draft materials such as a detailed meeting agenda and process to be used during the meeting; a template for performance tasks, including metadata to be collected; a draft scoring rubric which maintains comparability to Georgia Milestones and aligns with Georgia Milestones Achievement Level Descriptors (ALDs); and sample performance tasks for each grade or grade band which will serve as exemplars during the meeting.

The workshop will be a two-day event led by the external vendor-partner. NWEA content development and psychometric staff will be in attendance as observers of the process and owners of the project. There will be a breakout room for each grade 3-8 (for a total of six rooms each with one participant from each district) that will cover both ELA and Mathematics, with the main focus on assessing Writing. There will be approximately sixty Georgia educators involved in this workshop. The work will help train educators and increase their knowledge and confidence in the processes that will be used during performance task writing, construction, and scoring. We will also leverage the expertise of Georgia educators to ensure that the framework for performance tasks used meets the rigor of the Georgia Standards for excellence. NWEA will work with the GMAP districts to recruit educators for the meetings.

Following this meeting, NWEA and vendor-partners will move forward with building a plan to produce a sufficient number of performance tasks at each grade to account for attrition through the review and field testing to allow for complete coverage of the Georgia test blueprints. As mentioned earlier, the plan for coverage will be reviewed yearly to ensure a robust item bank for the GMAP assessment.

Application Selection Criteria	Required information from the SEA
<p>(a)(3) If the system will initially be administered in a subset of schools or LEAs in a State--</p> <p>(i) The strategies the SEA, including each SEA in a consortium, will use to scale the innovative assessment to all schools statewide, with a rationale for selecting those strategies;</p> <p>(ii) The strength of the SEA’s or consortium’s criteria that will be used to determine LEAs and schools that will initially participate and when to approve additional LEAs and schools, if applicable, to participate during the requested demonstration authority period; and</p> <p>(iii) The SEA’s plan, including each SEA in a consortium, for how it will ensure that, during the demonstration authority period, the inclusion of additional LEAs and schools continues to reflect high-quality and consistent implementation across demographically diverse LEAs and schools, or contributes to progress toward achieving such implementation across demographically diverse LEAs and schools, including diversity based on enrollment of subgroups of students describe in section 1111(c)(2) of the Act and student achievement. The plan must also include annual benchmarks toward achieving high-quality and consistent implementation across participating schools that are, as a group, demographically similar to the State as a whole during the demonstration authority period, using the demographics of</p>	<p>For the CCSD, GMAP and Putnam County models:</p> <ol style="list-style-type: none"> 1. Additional information regarding the scaling of the proposed innovative assessments to statewide use, specifically a clear description for how the State plans to choose just one of these innovative assessment options for final statewide implementation with sufficient time for the State to scale that one design for statewide implementation at the end of the IADA period (i.e., the State needs to implement its chosen design statewide in year five). 2. Information about benchmarks toward achieving implementation across participating schools that are, as a group, demographically similar to the State as a whole during the demonstration authority period for each model.

initially participating schools as a baseline.	
--	--

Cobb County School District

CCSD will expand to all schools in the district by the end of the second year of the pilot. CCSD has already implemented the assessment platform across 112 schools. As a whole, the district is demographically similar to the state of Georgia. CCSD will scale the assessments to other districts in the state in collaboration with the state.

Georgia MAP Assessment Partnership

During the IADA pilot, the process for adding partners/districts to the GMAP Consortium will begin with the district filling out a one page “Request to join GMAP Consortium” form and will submit that form to the GMAP district designated point of contact who will in turn forward to the GMAP senior program manager for NWEA to review. During the next GMAP Consortium monthly call, the application will be reviewed by the partners/districts and receive feedback from NWEA on the feasibility of adding the district. The GMAP Consortium will then vote to either accept the district as an Affiliate or a Collaborative partner. To become a Collaborative partner, a district should have administered MAP Growth for at least one year and agree to share their Georgia Milestone data with NWEA for comparability studies. The acceptance by the GMAP Consortium would then be forwarded to the GaDOE for approval to officially add the additional partner/district.

Providing a quality implementation and professional development plan is essential to the success of any assessment program. To ensure the sustained success of our proposed solution, NWEA has dedicated a senior program manager to the implementation of GMAP. This senior program manager serves as a primary point of contact and will oversee all aspects of the implementation plan ranging from initial training to data delivery.

NWEA is experienced in large state and district implementations of assessment, and has been training teachers, administrators, and districts to administer multiple interim assessments for 42 years. The same procedures used for onboarding in other states for state summative and state interim contracts and large district partners like Chicago Public Schools and Clark County Nevada is scalable for the state of Georgia, and as individual districts are added, district implementation procedures standard to the organization, including program management services, will be utilized.

Implementation of the through year model for all districts added begins with training test administrators on the product. An Administration course tailored to the GMAP solution will provide a detailed walk-through of how the test will be implemented and prepares proctors for test administration. This course will also provide insights into ways the test may be used in the classroom.

Prior to setting up the testing environment, the NWEA onboarding team works with each participating district to collect and document all data components necessary to build the testing environment (i.e., names of schools and identification of staff in key system administration roles, etc.). All information collected requires final sign-off by the district prior to the creation of testing events.

Once the testing environment has been created and tests loaded into the system, the implementation support specialist will verify that the environment includes the data components collected from the

district, and then conduct a “Go-Live” call with key staff. A verification of product functionality and a product walk-through is also performed during this call.

NWEA systems are able to scale to meet state needs. The current NWEA platform supports over sixty million student test events each year. The constraint engine and platform that will be used for this project is currently supporting summative assessments in Nebraska. Both systems are designed with highly scalable architecture and have been scaled successfully to support concurrent administrations, and MAP Growth supports more than 12 million students across taking tests in any given year across multiple administrations.

Putnam Consortium

The Putnam Consortium's original application detailed a scaling plan focused around three key points:

- Provide LEAs multiple entry points to participation, moving at a pace that allows schools to implement innovative assessments successfully
- Provide guardrails for participation and oversight for entry into participation
- Provide technology and capacity supports to LEAs during all levels of participation

The original application also detailed the representative nature of the districts that are currently using Navy with respect to demographic diversity of the state, as well as with respect to geographic and achievement diversity of the state.

At the time of the original application in December of 2018, 12 school districts were utilizing Navy. By February of 2019, three additional school districts contacted the Putnam Consortium and began using Navy to guide their instruction. The scaling patterns to this point, which show natural growth by word of mouth reaching diverse schools and districts, are evidence of the Putnam Consortium's promise for scaling statewide.

Currently at 8% of the school districts in Georgia, the Putnam Consortium will seek to increase to 10% of school districts by December of 2019 and then increase by 8-15% per year during the next 4 years of the pilot. This rate of increase would lead to 42-70% participation by Year 5 of the pilot, the time at which the state plans to select one system for complete statewide expansion. If selected, the Putnam Consortium will work alongside the state to implement its statewide expansion plan during the two extension years, as described below.

This rate of increase would also support the Putnam Consortium's plan to be finished with development and ready for external evaluation on an accelerated timeline, as a strong sample of diverse districts would be participating by Years 3 and 4.

To this point, the Putnam Consortium scaling has yielded a demographically representative sample of school districts. To ensure this continues to be the case as our consortium grows each year, leaders of the consortium and Navy Education will continue to support districts in all areas of the state to learn about joining the effort and team and to implement the assessment system successfully. Leaders of the consortium and Navy Education will seek to actively maintain diversity by monitoring the diversity of new members and, if needed, making efforts to recruit needed areas of diversity.

State of Georgia

For the five years of the demonstration authority, each of the three districts/consortia will work with stakeholders and technical experts to develop and implement innovative assessments within its district/consortium. Additional schools and districts not currently part of a consortium will have an opportunity to join a consortium during this time, as approved by the State Board of Education (SBOE) (per SB 362).

A report will be provided to the SBOE at least annually or upon request describing the progress made by all innovative assessment systems. At that time, all districts/consortia will have the opportunity to include additional school districts in the consortium for the following year. Districts that choose to join one of the approved innovative assessments would be required to seek amendment of their state performance contracts with the SBOE, which currently require them to use only Georgia Milestones. When the SBOE determines whether to approve the expansion of each innovative assessment to additional districts, they will consider demographic representation of the expanded consortium to ensure the consortium is demographically similar to the state as a whole.

The Georgia Department of Education (GaDOE) will use information from four sources to ultimately select one assessment system, approved by the State Board of Education, for possible statewide expansion.

First, GaDOE will utilize its superintendent, parent, and student advisory councils to collect stakeholder feedback on the innovative assessment pilot. The state will also develop an annual stakeholder feedback process to ensure all required stakeholders are kept informed of the innovative assessment pilot's progress and have the opportunity to provide feedback to inform development and implementation efforts. This work will be led by the Program Manager and Assessment Specialist described in Georgia's IADA application, who will oversee this demonstration authority, if awarded. Georgia will request funding for these positions from the state legislature during its 2020 session (if funded, the positions would be available for the 2020-2021 school year).

Second, based on the recommendation of the Governor, GaDOE will re-establish its Assessment Flexibility and Innovation Task Force, which will include 20 members – 10 appointed by the Governor and 10 appointed by the State School Superintendent. This task force will be charged with setting a vision for assessment in Georgia; providing feedback on the implementation of the innovative assessment pilot, including discussing stakeholder feedback and plans for scaling; and providing recommendations to the Governor and State School Superintendent regarding the innovative assessment pilot.

Third, GaDOE will utilize information and feedback from its technical assistance provider, who will provide technical assistance to participating consortia and produce an annual report of activities, needs, and next steps.

Finally, during Year 4 (January 2023), the GaDOE will request funding from the General Assembly to support the external technical evaluation required in Senate Bill 362 and described in Georgia's IADA application. If funded, GaDOE will release an RFP in Spring 2023 to seek services for the technical evaluation of all three innovative assessment systems, including assessing comparability with Georgia Milestones and content alignment studies. CCSD, GMAP, and Putnam will perform annual comparability studies prior to this technical evaluation, which will occur during Year 5 (2023-2024).

GaDOE will use information from the stakeholder feedback, task force, annual technical assistance, and external evaluation to select one assessment system for possible statewide expansion by the end of Year 5 (2023-2024). If one of these innovative assessment systems is selected for statewide expansion, Georgia will request a two-year extension of its demonstration authority. During the first year of the extension, GaDOE will plan for statewide implementation, which will occur during the second year of the extension. The year of planning will include securing the necessary contracts to be able to administer the assessment, orienting all districts to the new assessment system, and working with districts to ensure they are prepared to administer the new assessment during the second year of the extension (2025-2026).

If the state’s innovative assessment pilot progresses on a faster timeline than outlined in this section, GaDOE will seek to complete the technical evaluation on an accelerated timeline and work with the Assessment Flexibility and Innovation Task Force to provide recommendations to the Governor and State School Superintendent for the selection of an assessment system for possible statewide expansion.

Application Selection Criteria	Required information from the SEA
<p>(b)(1) The extent and depth of prior experience that the SEA, including each SEA in a consortium, and its LEAs have in developing and implementing the components of the innovative assessment system. An SEA may also describe the prior experience of any external partners that will be participating in or supporting its demonstration authority in implementing those components. In evaluating the extent and depth of prior experience, the Secretary considers --</p> <p>(i) The success and track record of efforts to implement innovative assessments or innovative assessment items aligned to the challenging State academic standards under section 1111(b)(1) of the Act in LEAs planning to participate; and</p> <p>(ii) the SEA’s or LEA’s development or use of--</p> <p>(A) Effective supports and appropriate accommodations consistent with 34 CFR part 200.6(b) and (f)(1)(i) and section 1111(b)(2)(B)(vii) of the Act for administering innovative assessments to all students, including English learners and children with disabilities, which must include professional development for school staff on providing such accommodations;</p> <p>(B) Effective and high-quality supports for school staff to implement innovative assessments and innovative assessment items, including professional development; and</p> <p>(C) Standardized and calibrated tools, rubrics,</p>	<p>For the CCSD model:</p> <ol style="list-style-type: none"> 1. Information regarding the adequacy of teacher training materials for performance task scoring (e.g., only one 45 minute training video was mentioned as a resource in the application). 2. Information regarding the qualifications of external psychometric consultants to be used on the pilot assessments. <p>For the GMAP model, information regarding the role of external organizations/partners in development of performance tasks for the pilot assessment and a description of their extent and depth of prior experience.</p>

<p>methods, or other strategies for scoring innovative assessments, with documented evidence of the validity, reliability, and comparability of annual summative determinations of achievement, consistent with 34 CFR part 200.105(b)(4) and (7).</p>	
--	--

Cobb County School District

The reference to training videos was intended to provide information about some of the resources that are currently available in CCSD for formative assessments. Training videos are not stand-alone resources, but instead are used by teachers as refreshers for face-to-face performance task scoring that is available. Information regarding plans for training scorers of the innovative assessment constructed response items can be found in the response to required information requested under application selection criteria (a)(2).

The CCSD has established qualification requirements for the psychometric partner(s) based on feedback from external consultants. The primary purpose of this position will be to manage and complete psychometric deliverables for CTLS-Assess. This include the design and development of technically sound psychometric structures, data collection, equating, calibration, linking, scaling, and psychometric analysis. This individual must possess the required psychometric skills and knowledge to support the district in developing sound, reliable, and valid assessments for an online administration. As such, this individual must have considerable knowledge of the principles and procedures used in test development, educational research, data analysis, as well as methods and designs for gathering, interpreting, and evaluating data and test results.

Georgia MAP Assessment Partnership

NWEA staff and partner organizations, such as the Georgia Center for Assessment (GCA), have rich experience in developing performance tasks and other assessment items. NWEA will utilize trained assessment professional item writers to create items and performance tasks for use in the assessment system and will engage the Georgia Center for Assessment to help facilitate the item review and scoring processes.

Specifically, GCA will:

- Host face-to-face meetings with educators such as focus groups, advisory boards, passage or item creation, and content reviews
- Facilitate educator meetings such as item content and bias reviews alongside NWEA staff
- Conduct hand-scoring activities for extended constructed response items used on GMAP

For all of these activities including hand-scoring, GCA has a 25-year history of conducting such activities within and beyond the state of Georgia.

Application Selection Criteria	Required information from the SEA
(b)(2) The extent and depth of SEA, including each SEA in a consortium, and LEA capacity to	For the CCSD model, more information regarding the strategies it is using, or will use, to mitigate

<p>implement the innovative assessment system considering the availability of technological infrastructure; State and local Ias; dedicated and sufficient staff, expertise, and resources; and other relevant factors. An SEA or consortium may also describe how it plans to enhance its capacity by collaborating with external partners that will be participating in or supporting its demonstration authority. In evaluating the extent and depth of capacity, the Secretary considers--</p> <p>(i) The SEA’s analysis of how capacity influenced the success of prior efforts to develop and implement innovative assessments or innovative assessment items; and</p> <p>(ii) The strategies the SEA is using, or will use, to mitigate risks, including those identified in its analysis, and support successful implementation of the innovative assessment.</p>	<p>risks to support successful implementation of the innovative assessment.</p>
--	---

Cobb County School District

The following strategies have been implemented to mitigate risks to support successful implementation of the innovative assessment system:

To mitigate risks associated with technological infrastructure, the CCSD Technology Services team will continue to support CTLS-Assess. Currently, the Technology Services departments supporting CTLS-Assess consists of 109 team members. The Technology Department has handled over 2,338 work requests in the past 4 years for CTLS-Assess.

The departments for Technical Support consist of the Customer Care Center/Help Desk, Network Operations, and Student Information Systems. The Customer Care Center/Help Desk teams include 8 support staff with varying levels of expertise that can troubleshoot calls from teachers and administrators and escalate to other technical staff/departments. The Network Operations team includes 12 support staff with varying levels of expertise and is the highest level of support prior to escalating to vendor support. The Student Information Systems team includes 10 support staff with varying levels of expertise and interfaces with vendors and for data and reporting services.

The Field Services Department consists of 48 Field Technicians and 4 Engineers. The Field Technicians work in assigned schools to support hardware and technical issues related to equipment and software/Web 2.0 and work collaboratively with the Instructional Technology team to provide technical support. The Engineers support CTLS-Assess for Life Cycle Management through software packaging and deployment.

The Instructional Technology Department consists of 27 Technology Training/Integration Specialist (TTIS). The TTIS team meets regularly with the Assessment Department for collaboration on CTLS-Assess. The TTIS team supports the development of training materials for the CTLS-Assess platform (slide decks, online resources and videos), partners with the

Assessment Department to provide ongoing training for teachers and leaders in the use of the CTLS-Assess platform, provides modeling/coaching to teachers in utilizing the assessment tools, and supports teachers with providing ongoing feedback on the CTLS-Assess platform for improvement or enhancements.

Additional risks associated with test security and fidelity to administration requirements exist. To mitigate these risks, a handbook specific to the innovative assessments will be developed. The handbook will include detailed information regarding processes and procedures for before, during, and after the assessment, including test security. The contents of the handbook will be consistent with the Georgia Student Assessment Handbook and the GaPSC Code of Ethics for Educators. Training will be provided to all Testing Coordinators by the CCSD Assessment Department. Testing Coordinators will redeliver the training to all test examiners. The current structure of support will be utilized to mitigate risks associated with fidelity to administration requirements. Personnel within the structure of support is as follows: CCSD Assessment Director, CCSD Assessment Supervisors, CCSD Program Managers, CCSD Technology Directors, and local school testing coordinators.

In addition, the assessment administration platform has been designed to support a secure assessment administration environment through the use of specific administration window that are set by select district level personnel; teachers do not have access to any secure test items or secure assessments; school leadership and district assessment personnel have the ability to monitor, in real time, which students are participating in each assessment, by teacher; assessment data reports do not disclose any item information, other than standard alignment and DOK level. See “Security feature of CTLS-Assess” on page 35 of Georgia’s original IADA application for additional technical security features.

To mitigate risks associated with validity and reliability, assessment blueprints will be refined to include content weights that are appropriate relative to CCSD instructional frameworks and Georgia Milestones Assessment System blueprints; blueprints include items that aligned to content standards, at the appropriate level of rigor for the standard; and are deemed reliable based on psychometric analysis; constructed response items will be scored using rubrics that have been in place in CCSD for three years, and a blind scoring method (described in a (a)(2)).

Application Selection Criteria	Required information from the SEA
<p>(c)(1) The extent to which the timeline reasonably demonstrates that each SEA will implement the system statewide by the end of the requested demonstration authority period, including a description of--</p> <p>(i) The activities to occur in each year of the requested demonstration authority period;</p> <p>(ii) The parties responsible for each activity; and</p> <p>(iii) If applicable, how a consortium’s member SEAs will implement activities at different paces and how the consortium will implement interdependent activities, so long as each non-affiliate member SEA begins using the innovative assessment in the same school year consistent with 34 CFR part 200.104(b)(2);</p>	<p>For the CCSD, GMAP and Putnam County models, more information is needed regarding project timelines.</p> <p>With regard to the CCSD timeline:</p> <ol style="list-style-type: none"> 1. Identify the parties are that will be performing or implementing each activity. 2. Clarify what the difference is between the “field tests” and “districtwide implementation at all the grade levels” that are both listed in Year 2 of the timeline. 3. Clarify what grade levels/groups of students will be in the field tests each year. 4. Indicate when training of teachers will occur prior to districtwide implementation.

	<p>With regard to the GMAP timeline:</p> <ol style="list-style-type: none">1. Clarify who would be developing test items. (If local teachers were to be involved in item development, please indicate when item writing training would occur and who would be provide the training.)2. Reconcile the development of performance tasks beginning in year 3, which is also the year the comparability study was to be conducted for R/LA and math. <p>With regard to the Putnam County timeline:</p> <ol style="list-style-type: none">1. Clarify what teachers and students will be doing each year regarding the pilot assessments2. Clarify when training for teachers would be provided in relation to the pilot assessment activities each year.
--	---

Cobb County School District

CCSD/GaDOE/External psychometric partner will be performing the activities.

The exact grade levels and groups of students for the field tests will be determined in consultation with our external psychometric partners during the first year of the demonstration authority.

Training of teachers on CTLS Assess will begin in July 2019, the week preceding the start of school on August 1, 2019.

Georgia MAP Assessment Partnership

Assessment items and performance tasks will be created continuously throughout the GMAP program to ensure blueprint coverage and best practices in item refresh rates are maintained. Summative Performance tasks will be used in year 3 and beyond to ensure blueprint coverage. All items used in the summative score will be created and scored by assessment professionals in a standardized manner.

Trained assessment professionals will be included in all phases of this program, including item development and scoring. Building assessments takes yearly planning to ensure long-term sufficiency of the content needed for a highly discriminating assessment. NWEA currently does this for the interim and summative assessments following guidelines which represent industry best practice. Using the full summative assessment blueprint for each grade and subject, NWEA will, at the beginning of each year, conduct an analysis of all existing content for GMAP through-year summative assessment to evaluate content coverage and identify development needs. Stimuli and items will be developed by professional content developers who are familiar with the program. Educators will be used in review activities to determine that content developed is appropriate for their students prior to field testing.

Content exists in multiple phases since each item will be field tested prior to operational usage. The content analysis will address all types of items, including performance tasks, for full blueprint coverage across each standard. In addition to new content development to fill blueprint gaps, NWEA will review its existing item bank of over 30,000 items to determine which items are aligned to Georgia Standards of Excellence for both interim and summative purposes through an alignment study at the end of year 1. Where deficiencies are found, NWEA staff will continue to develop content using professional content developers and review content with Georgia educators for use in the summative component of the GMAP assessments. The result will be a fully sculpted item pool that robustly meets the Georgia assessment blueprints and provides an equivalent testing experience for all students.

Performance Tasks are an important part of the through-year solution to elicit authentic work from students. This will be a key component of the solution to assessing writing from the ELA construct. NWEA is committed to producing performance tasks and have staff in both content development and psychometrics with experience in the area. However, we will also be engaging with in-state and national vendor-partners with additional expertise as discussed in section (b)(1).

Putnam Consortium

Teacher and Student Activities

Teacher roles include completing training to understand how to administer Navvy assessments with fidelity and security, how to interpret results appropriately, and how to utilize the instructionally-relevant feedback to support student learning in the classroom; administering the Navvy assessments; utilizing feedback from the Navvy assessments to inform instructional decisions; and facilitating student use of the platform to complete assessments and self-monitor learning.

Student roles include completing the Navvy assessments, learning how to interpret and utilize results from the Navvy assessments, and using Navvy results to self-monitor learning and to gain the support needed to learn.

Each year, teachers in participating districts will administer Navvy assessments to students in grades 3-8 for English and math, students in grades 5 and 8 science, high school students enrolled in a math or English course with a corresponding statewide assessment (2 courses per subject), and high school students enrolled in the science course with a corresponding statewide assessment.

In addition to the Navvy assessments, during the initial year(s) of the pilot, teachers will administer the Georgia Milestones to students in a representative sample of schools once per grade band in English language arts, math, and science to establish comparability. This representative sample of schools will require double-testing of a sample of students to establish comparability using the method described under 34 CFR 200.105(b)(4)(i)(B). Once comparability is established, students will no longer be required to complete the Georgia Milestones assessments.

Table 3 illustrates the grade and subject combinations where, for a representative sample of schools, both the innovative assessment system (Navvy) and the statewide academic assessments (Georgia Milestones) will be implemented in order to establish comparability: grade 3 (math); 4 (English language arts); 5 (science); 6 (math); 7 (English language arts); 8 (science); and in high school for the first course in math (Algebra I/Coordinate Algebra), for the second course in ELA (American Literature and Composition), and a course in science (Biology).

Grade	Math	ELA	Science
3	Navy & Milestones*	Navy	Local Assessments
4	Navy	Navy & Milestones	Local Assessments
5	Navy	Navy	Navy & Milestones
6	Navy & Milestones	Navy	Local Assessments
7	Navy	Navy & Milestones	Local Assessments
8	Navy	Navy	Navy & Milestones
High School Course 1	Navy & Milestones [Algebra I/Coordinate Algebra]	Navy [9 th Grade Literature and Composition]	Navy & Milestones [Biology]
High School Course 2	Navy [Geometry/Analytic Geometry]	Navy & Milestones [American Literature & Composition]	Local Assessments [Physical Science]

Table 3. *Assessment plan by grade and subject for establishing comparability*

*Navy is the innovative through-year assessment system and Milestones is the current end-of-year statewide assessment system.

Table 4 illustrates the grade and subject combinations where the innovative assessment system (Navy) will continue to be administered after comparability with the Georgia Milestones has been established.

Grade	Math	ELA	Science
3	Navy	Navy	Local Assessments
4	Navy	Navy	Local Assessments
5	Navy	Navy	Navy & Milestones
6	Navy	Navy	Local Assessments
7	Navy	Navy	Local Assessments
8	Navy	Navy	Navy

High School Course 1	Navy [Algebra I/Coordinate Algebra]	Navy [9 th Grade Literature and Composition]	Navy [Biology]
High School Course 2	Navy [Geometry/Analytic Geometry]	Navy & Milestones [American Literature & Composition]	Local Assessments [Physical Science]

Table 4. *Navy assessment plan by grade and subject*

Timeline for Teacher Trainings

Putnam County, the Putnam Consortium Innovative Assessment Executive Team, and Navy Education will seek to provide additional supports for teachers in three primary ways: (1) by holding quarterly innovative assessment summits, (2) by partnering with professional development experts who have had immense success in training educational leaders, and (3) by creating accessible, web-based training content to facilitate full-scale implementation supports for all participating administrators and teachers. Excerpts from the original application describing each type of training are provided below and followed by requested additional information about the timeline for the trainings.

Quarterly Innovative Assessment Summits

Quarterly summits provide the organizational form for participating LEAs to learn more about Navy, receive hands-on training for Navy, and participate in shared decision making for Navy. The summits will provide in-person training sessions for LEAs at various levels of participation, including (1) *Introduction to Navy*, an presentation providing information for districts interested in learning more about Navy assessments and the corresponding accountability system; (2) *On-boarding for Navy*, a training for new members on implementing the system with fidelity and security; (3) *Content Collaboratives*, meetings for in-person collaboration, discussion, and/or professional development about assessment content; (4) *Implementation Collaboratives*, meetings for collaboration, discussion, and/or professional development around implementation opportunities and challenges. These break-out groups through in-person meetings will facilitate collaboration among districts and feedback for continuous improvement.

These summits will take place during the fall (2 summits) and spring (2 summits) each year of the innovative pilot. A sample of teachers from each school district will participate in cohorts, attending all 4 summits in a given year. In subsequent years, new cohorts of teachers, as selected by the district, will participate in the summits.

Partnering with Professional Development Experts

The Putnam Consortium Innovative Assessment Executive Team and Navy Education will partner with the Institute for Performance Improvement (the “Institute”) to provide school leaders with training that will effectively support implementing Navy with fidelity and success. The Institute is currently partnering with the GaDOE to train and assess state and regional education center staff assigned to support district and school improvement, and, under ESSA, will train and assess principal supervisors and other district leaders, school leadership teams and academic coaches supported and served by the

GaDOE and the Regional Education Service Agencies, with state and regional staff serving as peer coaches (described further in the original application). For Navvy, the Institute will develop a training designed to support district level personnel to implement Navvy successfully using their evidence-based approach that ensures that participants can demonstrate proficiency, or they are provided feedback to determine what they have to do to reach performance.

This training has proven to be successful running through the school year, beginning in the summer, continuing through the improvement cycle, and along the year of teaching, assessing, and evaluating. The training will include 3 days before school starts and two more days in the Fall. Then, the training will include two days in the Winter, plus a virtual course, and will close out with two days in the early Spring. This training also uses a cohort model, with new cohorts being selected each year and the potential for multiple cohorts to be run during the same year.

Accessible Content for Full-scale Support

Online modules will be created to provide the necessary information for a district to launch Navvy. Content for online modules will be created by Navvy Education and members of the Putnam Consortium Innovative Assessment Leadership Team and online modules will be produced by a communications vendor.

These online modules will be available on-demand for all teachers and will be required for teachers to complete prior to administering Navvy assessments. The on-boarding modules provide an introduction to the Navvy assessments and show teachers how to administer the assessments with fidelity and security and how to appropriately interpret and use results from Navvy. Teachers in new member districts (in year 1 of participation) will complete online modules prior to the start of the new school year or during the first few weeks of school, as determined by the member district.

Application Selection Criteria	Required information from the SEA
<p>(c)(2) The adequacy of the project budget for the duration of the requested demonstration authority period, including Federal, State, local, and non-public sources of funds to support and sustain, as applicable, the activities in the timeline under paragraph (c)(1) of this section, including--</p> <p>(i) How the budget will be sufficient to meet the expected costs at each phase of the SEA’s planned expansion of its innovative assessment system; and</p> <p>(ii) The degree to which funding in the project budget is contingent upon future appropriations at the State or local level or additional commitments from non-public sources of funds.</p>	<p>With regard to the GMAP budget:</p> <ol style="list-style-type: none"> 1. Provide estimated costs for the work proposed. 2. Provide a menu of estimated costs of the possible deliverables described in the application. <p>With regard to the Putnam County budget:</p> <ol style="list-style-type: none"> 1. Describe the total expected costs for the work described in the application. 2. Provide a description of how the cost sharing between LEAs and the external vendors would be determined.

Georgia MAP Assessment Partnership

Costs below represent the base costs for development of a solution that can ultimately be scaled to meet the needs of a state assessment as well as the implementation over the 5-year pilot with

participating GMAP districts and does not reflect the additional costs for expanding to additional districts or scaling up statewide.

To minimize additional costs being passed onto GMAP districts, NWEA, as a non-profit, is seeking external funding from granting agencies to help defray the costs of design of the assessment and is putting up resources for much of the development work “in-kind”. In the first two-years of the pilot, GMAP districts are continuing to offer MAP Growth and to pay for their existing MAP Growth licenses using existing resources. Pricing for the pilot and operational years is still being finalized and is expected to be similar to the Georgia Milestones per-student assessment cost, allowing the assessment to be scaled statewide within existing state budget structures.

Table 5: GMAP Budget

NWEA Georgia Through Year Pilot Budget (development plus operational costs) FY19 - FY23	
Content Review, Item Development, Staffing & Workshops	\$ 3,525,000
Program Management, Support, & Research Services	\$ 2,525,000
Psychometrics and Data Analysis	\$ 1,967,500
Hand Scoring	\$ 1,860,000
Professional Learning	\$ 880,000
Alignment Studies	\$ 375,000
Standard Setting	\$ 125,000
Total	\$ 11,257,500

Putnam Consortium

The budget for the all five years of supporting the implementation and scaling of the Navvy assessment system is presented below. The budget below reflects the funds required to operate the IADA for the next 5 years and includes support for LEA implementation of an innovative assessment system and the collection of evidence to support the validity of the use of the assessment system for statewide accountability.

Budget items fall into three broad categories, each described below in the “Adequacy of the Budget” section:

- Supports for participation and implementation
- Statewide reporting
- Technical quality and comparability

Table 6: Putnam Consortium Budget

Category	Item	Year 1	Year 2	Year 3	Year 4	Year 5
Supports for participation and implementation	Quarterly Innovative Assessment Summit	\$70,000	\$70,000	\$70,000	\$70,000	\$70,000
	Stakeholder representatives input – Meeting with community group and professional organizations to provide input for continuous improvement	\$10,000	\$5,000		\$5,000	
	Communication --Materials for Stakeholders including parents and community members	\$5,000	\$5,000			
	Communication – video-based instructional videos for implementation support	\$10,000	\$8,000		\$5,000	
	Mini-grants for professional development for school leaders and teachers	\$20,000	\$20,000	\$20,000	\$30,000	\$40,000
	LEA Project Management Support Personnel	\$120,000	\$120,000	\$120,000	\$120,000	\$120,000
Statewide Reporting	Design and development of reports for annual summative determinations that are comparable with Milestones reports	\$20,000				
	Design and development of subgroup reporting for state and federal purposes	\$20,000				
	Providing reports in languages Milestones reports are provided in	\$80,000				
Technical Quality and Comparability	Potentially embed Milestones items into platform	\$15,000				
	Analyses for determining annual summative determinations	\$40,000	\$15,000			
	Annual Data Review	\$175,000	\$175,000	\$75,000	\$75,000	\$75,000
	Analyses for reporting subgroup performance	\$25,000				
	Analyses for assessing comparability between the two assessment systems	\$50,000	\$25,000	\$25,000		
	Setting standards for annual summative determinations	\$70,000	\$20,000			
	Independent Formative Evaluation	\$225,000				
	Technical report writing and documentation	\$25,000	\$15,000	\$15,000	\$15,000	\$15,000

Budget Clarifications: Cost Sharing

The costs for the assessment system are shared by Navvy Education and affiliate and participating LEAs. This budget does not include funding for building or development of the Navvy assessment system itself, as Navvy Education, LLC develops the system and will pay for all development of the system. Funds are required to support the state’s use of the system for accountability. Budgeted funds will be used to ensure district staff have proper training to implement the system with fidelity and success and to produce evidence of the technical merit and equity of the system for statewide accountability.

The budget describes two primary costs that affiliate and participating LEAs have associated with the implementation and scaling of the assessment system: (1) costs for district leaders and teachers to participate in training, (2) costs for district leaders and teachers to participate in collaborative leadership (i.e., attend quarterly summits, participate on conference calls, pay for travel associated with teacher representatives on review and planning meetings, contribute to shared communication materials). The Putnam Consortium will request funding from the state legislature to support (1) the collection of evidence to support the validity of the use of the assessment system for statewide accountability, (2) support for teacher trainings through mini-grants, and (3) support for LEA project management.

Application Selection Criteria	Required information from the SEA
<p>(d)(4) If the system includes assessment items that are locally developed or locally scored, the strategies and safeguards (e.g., test blueprints, item and task specifications, rubrics, scoring tools, documentation of quality control procedures, inter-rater reliability checks, audit plans) the SEA or consortium has developed, or plans to develop, to validly and reliably score such items, including how the strategies engage and support teachers and other staff in designing, developing, implementing, and validly and reliably scoring high-quality assessments; how the safeguards are sufficient to ensure unbiased, objective scoring of assessment items; and how the SEA will use effective professional development to aid in these efforts.</p>	<p>For the CCSD and GMAP models, more information regarding assessment items that are locally developed or locally scored, specifically:</p> <ol style="list-style-type: none"> 1. The strategies and safeguards (e.g., test blueprints, item and task specifications, rubrics, scoring tools, documentation of quality control procedures, inter-rater reliability checks, audit plans) the model has developed, or plans to develop, to validly and reliably score such items, including how the strategies engage and support teachers and other staff in designing, developing, implementing, and validly and reliably scoring high-quality assessments. 2. How the safeguards are sufficient to ensure unbiased, objective scoring of assessment items; and how the models will use effective professional development to aid in these efforts.

Cobb County School District

See the response to required information requested under application selection criteria (a)(2) and (b)(2).

Georgia MAP Assessment Partnership

All items used for the summative assessment will be scored by a professional scoring vendor using raters/scorers that are trained and monitored to ensure valid and reliable scores. No items used for the summative score will be locally created or scored.

Performance Tasks are an important part of the through-year solution to elicit authentic work from students. This will be a key component of the solution to assessing writing from the ELA construct. NWEA is committed to producing performance tasks and have staff in both content development and psychometrics with experience in the area. However, the consortia will also be engaging with in-state and national vendor-partners with additional expertise to help build a knowledge base for GMAP teachers and to ensure that development efforts are well aligned with Georgia Standards for Excellence expectations.

In 2019, GMAP is planning a joint workshop with external vendor-partners who specialize in hand scoring and performance task development to engage with educators in the state. The objectives of this workshop are to:

- Engage with educator stakeholders in Georgia to begin understanding how performance tasks are currently being used in their classrooms and how they may expand such usage with GMAP in a formative manner, and how this can inform the summative development efforts.
- In collaboration with Georgia educators, build the rubrics necessary to start development of performance tasks with content development staff who have expertise in this area.
- Increase capacity to produce performance tasks, both within NWEA and the GMAP consortium.

Prior to the workshop, NWEA and external vendor-partner content and psychometric staff will meet and collaborate to finalize an agenda and process, including roles and responsibilities and precise outcomes, to be used for and gained from the meeting. Prework will include the development of draft materials such as a detailed meeting agenda and process to be used during the meeting; a template for performance tasks, including metadata to be collected; a draft scoring rubric which maintains comparability to Georgia Milestones and aligns with Georgia Milestones Achievement Level Descriptors (ALDs); and sample performance tasks for each grade or grade band which will serve as exemplars during the meeting.

The workshop will be a two-day event led by the external vendor-partner. NWEA content development and psychometric staff will be in attendance as observers of the process and owners of the project. There will be a breakout room for each grade 3-8 (for a total of six rooms each with one participant from each district) that will cover both ELA and Mathematics, with the focus on assessing Writing. There will be approximately sixty Georgia educators involved in this workshop. NWEA will work with the GMAP districts to recruit educators for the meetings.

Following this meeting, NWEA and vendor-partners will move forward with building a plan to produce a sufficient number of performance tasks at each grade to account for attrition through the review and field testing to allow for complete coverage of the Georgia test blueprints. As mentioned earlier, the plan for coverage will be reviewed yearly to ensure a robust item bank for the GMAP assessment.

Summative performance tasks will be generated by professional content developers and there will be educator reviews to approve the tasks for field testing.

Appendix A

Cobb County School District

IADA Blueprint Prototype
5th Grade English Language Arts

The IADA Blueprint Prototype below represents the totality of all quarterly ELA assessments for 5th grade. Separate blueprints will be developed by teams of teacher leaders for each quarterly assessment. District personnel will review all blueprints prior to finalizing them for use with assessment development teams.

Domain/Reporting Category	Standards Assessed	Approximate % of Tests	Item Types
Reading & Vocabulary	ELAGSE5.RI (1, 2, 3, 4, 5, 6, 7, 8, 9)	53%	Standard Multiple Choice, Technology Enhanced
	ELAGSE5.RL (1, 2, 3, 4, 5, 6, 7, 9)		
	ELAGSE5.L (4, 4a, 4b, 4c, 5, 5a, 5b, 5c)		
Writing & Language	ELAGSE5.W (1, 1a, 1b, 1c, 1d, 2, 2a, 2b, 2c, 2d, 2e, 3, 3a, 3b, 3c, 3d, 3e, 4, 7, 8, 9, 9a, 9b)	47%	Standard Multiple Choice, Technology Enhanced, Constructed or *Extended Response (2 nd & 3 rd Quarter)
	ELAGSE5.L (1, 1a, 1b, 1c, 1d, 1e, 2, 2a, 2b, 2c, 2d, 2e, 3, 3a)		

*Extended Response items will be narrative, opinion, or informational/explanatory genre.

IADA Blueprint Prototype
5th Grade Mathematics

The IADA Blueprint Prototype below represents the totality of all quarterly mathematics assessments for 5th grade. Separate blueprints will be developed by teams of teacher leaders for each quarterly assessment. District personnel will review all blueprints prior to finalizing them for use with assessment development teams. In addition to the content standards included below, items will also be aligned to the appropriate the Standards of Mathematical Practice.

Domain/Reporting Category	Standards Assessed	Approximate % of Tests	Item Types
Operations and Algebraic Thinking	MGSE5.OA.1	10%	Standard Multiple Choice, Technology Enhanced
	MGSE5.OA.2		
	MGSE5.OA.3		
Number and Operations in Base 10	MGSE5.NBT.1	25%	Standard Multiple Choice, Technology Enhanced
	MGSE5.NBT.2		
	MGSE5.NBT.3 (a, b)		
	MGSE5.NBT.4		
	MGSE5.NBT.5		
	MGSE5.NBT.6		
	MGSE5.NBT.7		
Number and Operations - Fractions	MGSE5.NF.1	30%	Standard Multiple Choice, Technology Enhanced
	MGSE5.NF.2		
	MGSE5.NF.3		
	MGSE5.NF.4 (a, b)		
	MGSE5.NF.5 (a, b)		
	MGSE5.NF.6		
	MGSE5.NF.7 (a, b, c)		
Measurement and Data	MGSE5.MD.1	20%	Standard Multiple Choice, Technology Enhanced
	MGSE5.MD.2		
	MGSE5.MD.3 (a, b)		
	MGSE5.MD.4		
	MGSE5.MD.5 (a, b, c)		
Geometry	MGSE5.G.1	15%	Standard Multiple Choice, Technology Enhanced
	MGSE5.G.2		
	MGSE5.G.3		
	MGSE5.G.4		

IADA Blueprint Prototype
8th Grade English Language Arts

The IADA Blueprint Prototype below represents the totality of all quarterly ELA assessments for 8th grade. Separate blueprints will be developed by teams of teacher leaders for each quarterly assessment. District personnel will review all blueprints prior to finalizing them for use with assessment development teams.

Domain/Reporting Category	Standards Assessed	Approximate % of Tests	Item Types
Reading & Vocabulary	ELAGSE8.RI (1, 2, 3, 4, 5, 6, 7, 8, 9)	53%	Standard Multiple Choice, Technology Enhanced
	ELAGSE8.RL (1, 2, 3, 4, 5, 6, 9)		
	ELAGSE8.L (4, 4a, 4b, 4c, 5, 5a, 5b, 5c)		
Writing & Language	ELAGSE8.W (1, 1a, 1b, 1c, 1d, 1e, 2, 2a, 2b, 2c, 2d, 2e, 2f, 3, 3a, 3b, 3c, 3d, 3e, 4, 7, 8, 9, 9a, 9b)	47%	Standard Multiple Choice, Technology Enhanced, Constructed or *Extended Response (2 nd & 3 rd Quarter)
	ELAGSE8.L (1, 1a, 1b, 1c, 1d, 2, 2a, 2b, 2c, 3, 3a)		

*Extended Response items will be narrative, opinion, or informational/explanatory genre.

IADA Blueprint Prototype
8th Grade Mathematics

The IADA Blueprint Prototype below represents the totality of all quarterly mathematics assessments for 8th grade. Separate blueprints will be developed by teams of teacher leaders for each quarterly assessment. District personnel will review all blueprints prior to finalizing them for use with assessment development teams. In addition to the content standards included below, items will also be aligned to the appropriate the Standards of Mathematical Practice.

Domain/Reporting Category	Standards Assessed	Approximate % of Tests	Item Types
Numbers, Expressions, and Equations	MGSE8.EE.1	20%	Standard Multiple Choice, Technology Enhanced
	MGSE8.EE.2		
	MGSE8.EE.3		
	MGSE8.EE.4		
	MGSE8.NS.1		
	MGSE8.NS.2		
Algebra and Functions	MGSE8.EE.5	40%	Standard Multiple Choice, Technology Enhanced
	MGSE8.EE.6		
	MGSE8.EE.7 (a, b)		
	MGSE8.EE.8 (a, b, c)		
	MGSE8.F.1		
	MGSE8.F.2		
	MGSE8.F.3		
	MGSE8.F.4		
MGSE8.F.5			
Geometry	MGSE8.G.1	28%	Standard Multiple Choice, Technology Enhanced
	MGSE8.G.2		
	MGSE8.G.3		
	MGSE8.G.4		
	MGSE8.G.5		
	MGSE8.G.6		
	MGSE8.G.7		
	MGSE8.G.8		
	MGSE8.G.9		
Statistics and Probability	MGSE8.SP.1	12%	Standard Multiple Choice, Technology Enhanced
	MGSE8.SP.2		
	MGSE8.SP.3		
	MGSE8.SP.4 (a, b)		

IADA Blueprint Prototype
English Language Arts - High School American Literature

The IADA Blueprint Prototype below represents the totality of all quarterly ELA American Literature assessments for high school. Separate blueprints will be developed by teams of teacher leaders for each quarterly assessment. District personnel will review all blueprints prior to finalizing them for use with assessment development teams.

Domain/Reporting Category	Standards Assessed	Approximate % of Tests	Item Types
Reading & Vocabulary	RL.11–12 (1, 2, 3, 4, 5, 6, 9)	53%	Standard Multiple Choice, Technology Enhanced
	RI.11–12 (1, 2, 3, 4, 5, 6, 7, 8, 9)		
	L.11–12 (4, 4a, 4b, 4c, 5, 5a, 5b, 6)		
Writing & Language	W.11–12 (1, 1a, 1b, 1c, 1d, 1e, 2, 2a, 2b, 2c, 2d, 2e, 2f, 3, 3a, 3b, 3c, 3d, 3e, 4, 5, 7, 8, 9, 9a, 9b))	47%	Standard Multiple Choice, Technology Enhanced, Constructed or *Extended Response (2 nd & 3 rd Quarter)
	L.11–12 (1, 1a, 2, 2a, 2b, 3, 3a, 6)		

*Extended Response items will be narrative, opinion, or informational/explanatory genre.

**IADA Blueprint Prototype
High School Mathematics – Geometry**

The IADA Blueprint Prototype below represents the totality of all quarterly mathematics assessments for high school. Separate blueprints will be developed by teams of teacher leaders for each quarterly assessment. District personnel will review all blueprints prior to finalizing them for use with assessment development teams. In addition to the content standards included below, items will also be aligned to the appropriate the Standards of Mathematical Practice.

Domain/Reporting Category	Standards Assessed	Approximate % of Tests	Item Types
Congruence and Similarity	G-CO (01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13)	35%	Standard Multiple Choice, Technology Enhanced
	G-SRT (01, 01a, 01b, 02, 03, 04, 05)		
Circles	G-C (01, 02, 03, 04, 05)	15%	Standard Multiple Choice, Technology Enhanced
	G-GPE (01)		
Equations and Measurement	G-GPE (04, 05, 06, 07)	35%	Standard Multiple Choice, Technology Enhanced
	G-GMD (01, 01a, 01b, 02, 03, 04)		
	G-MG (01, 02, 03)		
	G-SRT (06, 07, 08)		
Statistics and Probability	S-CP (01, 02, 03, 04, 05, 06, 07)	15%	Standard Multiple Choice, Technology Enhanced