

GEORGIA INNOVATIVE ASSESSMENT PILOT PROGRAM

Please specify the end-of-grade and/or end-of-course assessments for which evidence is being provided for the innovative assessment.

ELA	MATHEMATICS	SCIENCE	SOCIAL STUDIES
<input type="checkbox"/> Grade 3	<input type="checkbox"/> Grade 3		
<input type="checkbox"/> Grade 4	<input type="checkbox"/> Grade 4		
<input type="checkbox"/> Grade 5	<input type="checkbox"/> Grade 5	<input type="checkbox"/> Grade 5	
<input type="checkbox"/> Grade 6	<input type="checkbox"/> Grade 6		
<input type="checkbox"/> Grade 7	<input type="checkbox"/> Grade 7		
<input type="checkbox"/> Grade 8	<input type="checkbox"/> Grade 8	<input type="checkbox"/> Grade 8 <input type="checkbox"/> HS Physical Science (Grade 8)	<input type="checkbox"/> Grade 8
<input type="checkbox"/> American Literature and Composition	<input type="checkbox"/> Algebra I/Coordinate Algebra	<input type="checkbox"/> Biology	<input type="checkbox"/> U.S. History

For each of the assessments selected in the table above, evidence will need to be submitted for each of the criteria in the seven categories below (alignment and comparability, technical quality, accessibility and accommodations, test administration and security, stakeholder engagement, accountability, and conflict of interest). Note that all evidence submitted should be based on grade-level items only. Off-grade items can be included on assessments but cannot be included in the evidence required below.

1 ALIGNMENT & COMPARABILITY

Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents* (pages)	Commentary (Optional)
<p>1 Do you have an independent alignment study between the innovative assessment and the Georgia academic content standards (GSEs) for all grades, content areas, and courses?</p> <p>Note: The revised mathematics GSEs are expected to be operational for the 2022-2023 school year and the revised ELA GSEs are expected to be operational for the 2023-2024 school year.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Alignment study report	<Consortium A Alignment Report 2022.docx> (1-35)	
<p>2 Does the alignment study indicate that the innovative assessment adequately reflects Georgia academic content standards for all grades, content areas, and courses in terms of categorical concurrence, balance of representation, depth of knowledge, and range of knowledge?</p> <p>Note: If the innovative assessment is computer adaptive, documentation should demonstrate procedures that ensure the item pool and content constraints result in good alignment at the student level across all ability levels.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Alignment study report <ul style="list-style-type: none"> • Similar to alignment of Georgia Milestones Test blueprints indicating depth of knowledge ranges/cognitive complexity levels Item and passage specifications Item selection procedures	<Consortium A Alignment Report 2022.docx> (32-33)	
<p>3 Does the innovative assessment classify students into four achievement levels that are consistent (representing similar levels of knowledge and skill) with those reported for Georgia Milestones?</p>	<input type="checkbox"/>	<input type="checkbox"/>	Achievement level descriptors	<Consortium A Statewide Performance SY21-22.pdf> (2)	

	<p>Note: Direct adoption of Georgia’s ALDs is recommended to satisfy this criterion. If other ALDs are used, they must be justified and the alignment to the Georgia ALDs evaluated.</p>					
4	<p>Are summative classifications of students into the four achievement levels consistent between the innovative assessment and Georgia Milestones for all students and for all subgroups of students across all grades, content areas, and courses?</p> <p>Note: A standard setting is not expected, rather, empirical methods can be used to set cut scores on the innovative assessment that results in consistent student classifications into achievement levels. If the innovative assessment contains any off-grade level items, achievement level classification should be determined using only items that measure on-grade level standards (i.e., the grade in which the student is enrolled) and uses that determination for reporting and accountability. Consortia should also be aware that end-of-course assessments contribute 20% to course grades. The grade conversion score (GCS) is tied to the scale score cuts for Developing Learner and Proficient Learner. Specifically, for Georgia Milestones, the GCS ranges from 0 to 100. GCS=0 is set to the LOSS, GCS=100 is set to the HOSS. GCS=68, 80, and 92 are set to the scale cuts between achievement levels (1/2; 2/3; 3/4). A linear</p>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Classification consistency methods report, including achievement level classification consistency values and 4 x 4 contingency table for all grades, content areas, and courses for all students and all subgroups of students:</p> <ul style="list-style-type: none"> • Exact Agreement (>0.7) • Exact + Adjacent Agreement (>0.9) • Quadratic Weighted Kappa (>0.85) <p>The report or associated evidence should document, as applicable: methodology, calibration model(s), assumption check results, reliability, mean/range item difficulty, distribution of item types across the scale, student sample exclusions and impact of exclusions, consistency of results by demographic subgroups, comparability of administration conditions (e.g., speededness, format). The classification consistency report should also include an analysis of how comparable student grades are likely to be for end-of-course assessments given the GCS method.</p>	<p><Consortia A vs. Milestones Performance Level Classification Consistency (SY21-22).docx> (1-30; results pages 28-31)</p>	

	transformation is applied to obtain the GCS values between the points above.					
5	<p>Are the students who participate in the innovative assessment representative of the state in terms of demographic composition and achievement?</p> <p>Note: If the answer to this question is no, then provide evidence demonstrating how the sample has been weighted or adjusted to represent the state when necessary.</p>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Table of sample vs. state demographics and achievement (include all subgroups reported in Georgia for accountability)</p> <p>Description of weighting methods or other mechanisms for generalizing sample results to the state.</p>		
6	<p>Do you have a plan for conducting annual comparability analyses between the innovative assessment and Georgia Milestones throughout the remainder of the IADA period?</p> <p>Note: Comparability analyses will require double testing of Georgia Milestones and the innovative assessment for a sample of grades and subjects.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Comparability analysis plan		

*The Evidence Documents column can either contain the file name(s) of the relevant artifact(s), or a hyperlink to the document.

2 TECHNICAL QUALITY

Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents (pages)	Commentary (Optional)
1 Have you worked with experts to ensure technical quality, validity, reliability, and psychometric soundness of the innovative assessment?	<input type="checkbox"/>	<input type="checkbox"/>	CVs/qualifications of technical team Meeting agendas or meeting summaries (e.g., internal meetings, WestEd technical assistance meetings, TAC meeting transcripts, other consultant meetings)		

2	<p>Have you established reliability evidence for the summative scores, subscores, and achievement levels generated from the innovative assessment consistent with nationally-recognized testing standards?</p> <p>Notes: For preliminary or on-demand results/scores, demonstrate the technical evaluation procedures used to evaluate consistent reliability, including evaluation of model assumptions/parameters/scale stability. As a point of comparison, the majority of Georgia Milestones EOG and EOC assessments have reliability values of 0.9 and above. Include subscore reliability, but strict reliability criteria will not be required. Decision consistency and accuracy values should be similar to those reported for Georgia Milestones.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Reliability section of the technical report (include overall reliability, subscore reliability, conditional standard errors of measurement, decision consistency, and decision accuracy)		
3	<p>Have you established validity evidence for the innovative assessment consistent with nationally-recognized testing standards?</p> <p>Note: Much of the Comparability assurances criteria also provide validity evidence. Content evidence is most critical, relations to other variables will be available through comparison to Georgia Milestones, and validity evidence should be organized around the five sources of validity evidence described in <i>The Standards</i>. Evidence of test consequences, especially as it relates to the theory of action should be provided as soon as possible.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Validity section of the technical report Blueprints, test specifications, alignment studies		
4	Is the innovative assessment designed to assess student achievement based on	<input type="checkbox"/>	<input type="checkbox"/>	Score distributions		

	grade-level state academic content standards in terms of content and cognitive processes, including higher-order thinking skills, and to adequately measure summative student performance across the full performance continuum for all students, except students with the most significant cognitive disabilities?			Test blueprints, assessment guides, or other documents indicating depth of knowledge ranges Summary of item types Item and passage specifications Cognitive labs or other studies addressing student cognitive processes Analyses of test information functions demonstrating precision across the performance continuum or other demonstration of information function across the performance continuum CSEM across the scale/at the cut points Analyses (e.g., differential item functioning (DIF), differential test functioning (DTF) analyses) that identify possible bias or inconsistent interpretations of results across student groups Alignment studies		
5	Do you produce individual student score reports?	<input type="checkbox"/>	<input type="checkbox"/>	Example student report Score interpretation guide		
6	Do you produce aggregate score reports?	<input type="checkbox"/>	<input type="checkbox"/>	Example classroom, school, district, consortium reports Score interpretation guide		
7	Have you collected evidence that students, parents, educators, and school leaders are able to use your score reports to make valid score interpretations? Note: Include information about the representativeness of the sample for each stakeholder group.	<input type="checkbox"/>	<input type="checkbox"/>	Reports from cognitive labs, focus groups, etc.		
8	Are score reports provided in a timely manner?	<input type="checkbox"/>	<input type="checkbox"/>	Reporting timeline (e.g., number of days between the administration and when score users are provided with preliminary and/or final results along		

				with activities occurring between the two milestones)		
9	Have you incorporated principles of Universal Design for Learning into your innovative assessment?	<input type="checkbox"/>	<input type="checkbox"/>	Test development chapter of technical report Accessibility/UDL reports		
10	Have you developed a maintenance and evaluation plan to address longitudinal scale stability, identification and mitigation of parameter drift, and bank maintenance?	<input type="checkbox"/>	<input type="checkbox"/>	Psychometrics, research, and evaluation section of the technical report Details on item pool		

3 ACCESSIBILITY & ACCOMMODATIONS

All students who currently participate in Georgia Milestones must be able to participate in the innovative assessment in order to use the innovative assessment in lieu of Georgia Milestones. A crosswalk of accessibility and accommodation features available on Georgia Milestones and available on the innovative assessment should be provided such that it is possible to see at a glance whether all of the accessibility and accommodation features will be available, and if not, how students will be validly assessed using an alternative accessibility mechanism. Any differences in the way accessibility or accommodation features work in the innovative assessment as compared to Georgia Milestones should be indicated. Over time, the accessibility and accommodation features available for use on the innovative assessment should improve to reach industry best-practice.

Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents (pages)	Commentary (Optional)
1 In participating schools, are all students, except those with the most significant cognitive disabilities, participating in the innovative assessment?	<input type="checkbox"/>	<input type="checkbox"/>	Participation rate report Table of sample vs. state demographics and achievement		
2 Are students with disabilities provided with appropriate accommodations as defined by their IEP/IAP?	<input type="checkbox"/>	<input type="checkbox"/>	Relevant sections of the accommodations manual List of available accommodations Braille and VSL materials/resources Results of analyses and/or expert review indicating that accommodations do not alter the construct (e.g., classification consistency studies, DIF studies, person fit studies)		

3	Are English learners provided with appropriate accommodations as defined by their EL/TPC?	<input type="checkbox"/>	<input type="checkbox"/>	Relevant sections of the accommodations manual List of available accommodations Results of analyses and/or expert review indicating that accommodations do not alter the construct (e.g., classification consistency studies, DIF studies, person fit studies)		
4	Do all provided accessibility tools and accommodations comply with all federal laws, including, but not limited to, IDEA, ADA, Section 504 of the Rehabilitation Act of 1973, Title I, ESEA, and FERPA?			Relevant sections of the accommodations manual		

4 TEST ADMINISTRATION & SECURITY

If some of the test administrations do not contribute to a summative score, then the test administration and security requirements could be reduced. However, items from high-stakes administrations should not also be used during low-stakes administrations.

Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents (pages)	Commentary (Optional)
1 Has GOSA monitored your test administrations? Note: The consortia should work with GOSA and GaDOE to develop and implement a test monitoring plan.	<input type="checkbox"/>	<input type="checkbox"/>	Communications with GOSA GOSA audit reports		
2 Do you have policies and procedures to ensure standardized test administration?	<input type="checkbox"/>	<input type="checkbox"/>	Test coordinator manuals, test administration manuals, accommodations manuals, test preparation materials for students and parents, other documents provided to schools and teachers that address standardized test administration and		

				any accessibility tools and features available for the assessments Irregularity reports Proctor/test site training certificates		
3	Are all school staff that are involved in the test administration trained on standardized procedures and test security protocols?	<input type="checkbox"/>	<input type="checkbox"/>	Training presentation slides, documents, agendas Student assessment handbook Administration protocols Accessibility and accommodations manual Other comprehensive test administration policy documents Proctor/test site training certificates		
4	Do you have a process for monitoring the innovative assessment administration?	<input type="checkbox"/>	<input type="checkbox"/>	Relevant sections of the test coordinator manual Consortium monitoring analysis/report		
5	Do you have policies and procedures to prevent testing irregularities and ensure the integrity of test results?			Relevant sections of the student assessment handbook or assessment administration protocol manual Irregularity reports Monitoring results Data forensic methods and results		
6	Do you have test security policies and procedures to protect the integrity and confidentiality of test materials, test-related data, and personally identifiable information as established by the Family Education Rights and Privacy Act (FERPA) and the Georgia Student Data Privacy, Accessibility and Transparency Act of 2016?			Relevant sections of the student assessment handbook, test administration manual		

5 STAKEHOLDER ENGAGEMENT

Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents (pages)	Commentary (Optional)
----------	-----	----	-------------------------------	----------------------------	-----------------------

1	<p>Did you develop the innovative assessment in collaboration with stakeholders representing the interests of students with disabilities, English learners, and other vulnerable populations; teachers, principals, and other school leaders; parents; and civil rights organizations?</p> <p>Note: Consultation with these groups is required at the beginning on the project; ongoing consultation is not required.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Meeting schedules, meeting agendas, letters of support, meeting participants and associated demographics or background information		
2	<p>Did you develop capacity for educators and schools and districts leaders to implement the innovative assessment, interpret results, and communicate with stakeholders?</p>	<input type="checkbox"/>	<input type="checkbox"/>	Training agendas and presentations, meeting schedules, meeting agendas, other training materials, assessment guides, study/resource guides, item and scoring samplers, professional learning offerings, score interpretation guide, data on stakeholder participation in training for test administration, official logs for materials distribution, stakeholder survey results		

6 ACCOUNTABILITY

CCRPI growth, gaps, and literacy measures do not need to be strictly comparable, nor are the innovative assessments required to use the same methods that are currently used for Georgia Milestones. The methods do need to be justified and defensible.

Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents (pages)	Commentary (Optional)
<p>1 Do you have a process for identifying students uniquely within and across years so that students' assessment data, schools,</p>	<input type="checkbox"/>	<input type="checkbox"/>	Database with unique student identifiers (e.g., Georgia Testing Identifier [GTID])		

	<p>districts, demographic information, etc. can be used for accountability purposes?</p> <p>Note: The consortia should work with GaDOE to develop a data layout and reporting timeline.</p>					
2	<p>Is the percentage of students (overall and by subgroup) that you assessed in the current academic year at least as high as the percentage assessed using Georgia Milestones in the year previous to the start of the pilot (i.e., 2018-2019)?</p>	<input type="checkbox"/>	<input type="checkbox"/>	Participation rate report		
3	<p>Do you produce a single, summative score for every student?</p> <p>Note: If there is more than one administration during the academic year (e.g., a through-year model), specify which administrations contribute to the summative score and how scores are combined. This description should provide a clear rationale for the calculation of the summative score.</p>	<input type="checkbox"/>	<input type="checkbox"/>	Scoring section of the technical report		
4	<p>Do you produce a growth measure that can be used for the CCRPI Progress component?</p>	<input type="checkbox"/>	<input type="checkbox"/>	Growth measures section of the technical report		
5	<p>Do you produce an achievement measure that can be used for the CCRPI Content Mastery and Closing Gaps components (alignment to Beginning, Developing, Proficient, and Distinguished Learner achievement levels)?</p>			Scoring section of the technical report		
6	<p>Do you produce a literacy (Lexile) measure that can be used for the CCRPI Readiness component?</p>	<input type="checkbox"/>	<input type="checkbox"/>	Classification consistency methods report		

	Note: Classification consistency should be demonstrated for two designations: Reading Status as reported for Georgia Milestones and the literacy indicator as reported for CCRPI.					
7	Do you produce subgroup results consistent with federal accountability and reporting requirements (e.g., race/ethnicity, gender, English Learners, students with disabilities, migrant, homeless, foster, parent on active military duty, economically disadvantaged)?	<input type="checkbox"/>	<input type="checkbox"/>	Consortium summary report		

7 CONFLICT OF INTEREST

	Criteria	Yes	No	Examples of Relevant Evidence	Evidence Documents (pages)	Commentary (Optional)
1	Is there a conflict of interest (financial or otherwise) for the interested parties participating in the pilot program?	<input type="checkbox"/>	<input type="checkbox"/>	N/A	N/A	
2	Do all activities that are related to this pilot abide by local procurement requirements?	<input type="checkbox"/>	<input type="checkbox"/>	N/A	N/A	