



STATE OF HAWAII
DEPARTMENT OF EDUCATION

P.O. BOX 2360
HONOLULU, HAWAII 96804

OFFICE OF STRATEGY, INNOVATION AND PERFORMANCE

June 29, 2020

Frank T. Brogan
Assistant Secretary for Elementary and Secondary Education
U.S. Department of Education
Office of Elementary & Secondary Education
400 Maryland Ave. SW
Rm. 3E124
Washington, DC 20202

Dear Assistant Secretary Brogan:

The Hawaii Department of Education (HIDOE) has received and reviewed the U.S. Department of Education's (ED) letter dated April 28, 2020 declining to approve the HIDOE to implement its Innovative Assessment Demonstration Authority application. In accordance with the letter informing HIDOE of the opportunity to provide additional information, attached is HIDOE's response in Addendum 2.

The HIDOE theory of action for its innovative assessment design is to create a system that, pursuant to section 1111(b)(2)(B)(x) of the ESEA, "allow[s] parents, teachers, principals, and other school leaders to understand and address the specific academic needs of students, and that are provided to parents, teachers, and school leaders, as soon as is practicable after the assessment is given, in an understandable and uniform format, and to the extent practicable, in a language that parents can understand."

The HIDOE proposes the use of a shortened, summative computer adaptive test (CAT) that measures the full breadth and essentially the same depth of the Common Core State Standards (CCSSs) as the current statewide assessment. The shortened assessments cover the same targets/standards as the Smarter Balanced Assessments computer adaptive tests using approximately half of the number of items to produce an overall score. HIDOE acknowledges the elimination of the single writing performance task may reduce the capacity of the assessment to measure the depth of the writing standards, but HIDOE will include teacher-developed and reviewed writing as well as math performance tasks as common assessments as part of the Classroom Based Assessment System (CBAS). The data gathered from the CBAS

will assist educators in identifying areas of academic strengths and needs, relative to other metrics such as class grades and formative assessments.

Hawaii Comprehensive Assessment Program (HICAP) teachers will participate in professional development opportunities in order to deepen their understanding of high quality assessments aligned to the learning expectations and how to create performance tasks that are appropriate for the students in the classroom.

Therefore, HIDOE proposes a stepwise approach to creating writing and math performance tasks that may be used as common CBAS along with the shortened summative. Professional development and training will occur in Year 1 to assist teachers in creating and reviewing performance tasks, while relying on a shortened summative assessment to provide results to be used for accountability and reporting. HIDOE believes that the best way to increase teacher knowledge, participation and buy-in to the use of a common performance task is to ensure they have the proper understanding and support to develop, implement, and interpret these CBAS.

HIDOE looks forward to collaborating with the ED to build a truly innovative system that results in increased assessment literacy and academic achievement.

Sincerely,

Teri Ushijima, Ed.D.
Assessment and Accountability Director

TU:br

Attachments: Addendum 2: Additional Information or Revision in Hawaii's Innovative
Assessment Demonstration Authority Plan
Appendix Q: Quality Assurance and Assessment Monitoring Site Visit
Appendix T: Performance Assessment Training

c: Office of the Superintendent
Office of Strategy, Innovation, and Performance
Assessment and Accountability Branch

**Addendum 2: Additional Information/Revision in Hawaii’s Innovative Assessment Demonstration Authority Plan
Hawai’i Department of Education
June 29, 2020**

Regulatory Requirement	Required information from the SEA	HIDOE Response
<p><u>(b) Innovative assessment system.</u> 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</p>	<ul style="list-style-type: none"> • Evidence requested in sections (b)(2) and (b)(7) through (b)(9) below. 	

<p>any non-participating school within a participating LEA; and (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>(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 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</p>	<ul style="list-style-type: none"> ● Evidence that the proposed innovative assessment used for accountability purposes (the shortened summative assessment) is sufficiently aligned to the full depth of the State’s academic content standards, specifically: <ul style="list-style-type: none"> ○ A demonstration that the proposed innovative test blueprint proposed for accountability determinations assesses the same depth of the academic content standards 	<p>The Hawai‘i Department of Education (HIDOE) proposes the use of a shortened, summative, computer adaptive test (CAT) that measures the full breadth and essentially the same depth of the Common Core State Standards (CCSSs). The shortened assessments cover the same targets/standards as the Smarter Balanced Assessments Computer Adaptive Tests using approximately half of the number of items to produce an overall score. While HIDOE acknowledges the elimination of the single writing performance task may reduce the capacity of the shortened summative assessment to measure the depth of the writing standards, the simulation results of the shortened summative assessments indicate a 0.98 correlation with the full summative assessment. HIDOE maintains that the overall score generated for accountability purposes easily meets all technical</p>

<p>(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>as the statewide assessment, especially given the proposed differences in item types (e.g., no constructed response items) when compared to the statewide assessment.</p>	<p>quality requirements.</p> <p>To further address the full depth and breadth of content standards and higher order thinking skills, teachers participating in the pilot program will be required to administer at least one common performance task aligned to higher order thinking skills. Information from these assessments will supplement information gathered through the computer adaptive test described above. The development and administration of the performance assessments will be monitored by content and assessment specialists who will advise and support this work to ensure alignment to the CCSSs at the appropriate depth of knowledge levels. Teachers will be provided professional development in order to deepen understanding of high quality assessments aligned to the learning expectations and how to create an assessment appropriate for the students that are currently in the teacher’s class.</p> <p>Teacher-created and peer-reviewed classroom-embedded assessments, including performance-based tasks can address the depth of the CCSS in ways that are not possible with a single-opportunity computer-based summative assessment. Assessment tasks administered as part of instruction are better suited to measure deeper learning than are standardized performance tasks administered as part of an end-of-year on-line summative assessment.</p> <p>The CAT, built upon Smarter Balanced items, will provide a student’s overall scale score and proficiency level for accountability purposes. The results of the Classroom-Based Assessment System (CBAS) will populate the standards-based grading system</p>
---	--	--

		<p>immediately upon completion of a computer-based test or uploading of classroom-based assessment results by the teacher. Each student will receive a proficiency level (i.e., grade) for each grade level content standard. These proficiency levels will be compiled at the claim level (matching the Smarter Balanced target and claim categorizations) using a ‘preponderance of evidence’ and ‘most-recent evidence’ methodology. The CBAS claim level scores will be compared with those generated by the shortened summative CAT and as long as the CBAS score falls within 1.5 standard deviations of the CAT-generated claim level score the CBAS score will be reported on the end-of-year Family Report indicating whether or not the student achieved “Below Standard” or “Above Standard” with an indicator representing a standard error of measure. If there is disparity between the CBAS claim level score and the CAT-generated claim level score then “Not enough information” will be printed on the end-of-year family report. This will also be an indication to the project team that there is a need for further consultation and support for a particular teacher(s). There will be four Performance Assessment Institutes (see Appendix T) where teachers will learn about evidence-centered design aligned to state content standards using principles of Universal Design for Learning. Each institute will focus on aspects to develop teachers’ ability to create and score aligned assessments. At the end of each academic year, there will be a comparative analysis of the classroom assessment and the shortened summative assessment to further study whether the 1.5 standard deviation needs adjustments. Further professional development will be provided if deemed necessary.</p>
--	--	---

		<p>The CBAS system will aggregate results in real time for a teacher to review and make decisions about next steps for instruction that are specific for a group or individual students. The results of the shortened, summative CAT will also be aggregated and compared with the CBAS aggregate results. Teachers who have aligned their CBAS assessments with the CCSSs will find that the CBAS aggregations match those of the shortened, summative CAT aggregations. In cases where there is a mis-match between the two systems, the information can then be used to discuss alignment issues with the teacher and to suggest or require professional development the following school year.</p>
<p>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--</p> <p>(i) The student’s mastery of the challenging State academic standards under section 1111(b)(1) of the Act for the grade in which the student is enrolled; or</p> <p>(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>	<ul style="list-style-type: none"> ● Evidence that the shortened pilot assessment will sufficiently describe the student’s mastery of the State’s challenging academic standards, given that the innovative assessment is substantially different in length than the current statewide assessment, and does not contain constructed response items. Evidence provided to address requirement (b)(2) above may also address the concern for this critical element. 	<p>[Please see b(2)(i) response above.]</p> <p>EDs feedback appears to indicate a distinction between the innovative assessment and the annual summative assessment. These are two components of a single, integrated innovative assessment system. Yes, HIDOE argues that the modified (shortened) CAT summative assessment employed within the innovative assessment project meets all of the validity and reliability requirements to provide overall performance scores and identify students’ proficiency levels to meet the requirements of the accountability measures.</p> <p>Additionally, the classroom assessment component of Hawaii’s innovative assessment system provides finer-grained information than a single, end-of-year assessment about students’ knowledge and skills within the content domains as laid out in Smarter Balanced claims and targets. As the pilot progresses, there will be an expanding pool of shared classroom-based assessments that have been vetted and tested,</p>

		<p>such that it may be possible for shared teacher-developed classroom-administered performance assessment to be incorporated, along with CAT results, in the calculation of students' overall scale scores and performance levels.</p>
<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(e);</p>	<ul style="list-style-type: none"> • Evidence that the pilot assessments will provide disaggregated results by each subgroup of students, including timely data for teachers, principals and other school leaders, students, and parents, given that local assessment results will be reported in conjunction with State assessment results on parent reports (e.g., provide a rationale for reporting results of State assessments on the same document as local assessment results that are not standardized). 	<p>First, the shortened summative CAT provides reliable and valid results for individuals and by subgroups and schools.</p> <p>Additionally, the CBAS will provide classroom-based assessment results disaggregated by each subgroup of students. Teachers may use the results to determine potential needs for additional instruction by individual, subgroup and class.</p> <p>Teachers, students and parents will have access to individual results in real time. Using a secure login, parents and students may receive the results of completed assessments and review indicators of performance at the content standard level. This will provide an opportunity to make decisions about additional instructional needs or whether the student is ready for the next level of instruction. This timely feedback, made available in real time to those best positioned to make decisions about achievement, will streamline the learning process and reduce gaps in knowledge and skills.</p> <p>The shortened, summative CAT results will be made available to teachers immediately upon completion of the assessment. Results will be disaggregated by subgroups and provide indicators of achievement at the overall, claim and target levels for a group of students.</p> <p>The HIDOE Innovative Assessment Stakeholder</p>

	<p>group provided feedback indicating that the current assessments used for accountability purposes ‘do not embody what is being taught in classrooms.’ The stakeholders also expressed a desire to use ‘authentic assessments’ that teachers administer throughout the school year as part of the reporting of end-of-year assessment results. The HDOE proposes that both the shortened, summative CAT results and the validated CBAS results be included on the Family Report to honor the recommendations made by the stakeholder group and those expressed by Hawaii State Teacher Association members and educators throughout the state. The shortened, summative CAT overall score and proficiency level, along with revised achievement level descriptors, will be included in the Family Report and identified as the score used for state and federal accountability purposes. The CBAS claim level scores will also be included on the Family Reports and identified as being generated by teacher-created and administered assessments. The claim level scores will only be included on the Family Report if the score falls within a range of acceptability based on the comparability analyses conducted by HDOE and its technical consultants. If the CBAS claim score does not fall within the acceptable range then “Not enough information” will be printed on the Family Reports. Using this methodology to ‘verify’ the claim level scores, only scores that align with the summative results will be included in the Family Report as described by the Critical Elements for reporting.</p>
--	---

	<ul style="list-style-type: none"> Evidence that the local assessment component of the innovative pilot are ready to be implemented and combined with the short summative component in the 2020-21 school year to calculate the overall summative score and to be included as part of the accountability system for participating schools. 	<p>The HIDOE is in the process of securing a contract for a Classroom-Based Assessment System in the summer of 2020. An RFP was posted on February 5, 2020, and, after a brief hiatus due to the COVID-19 pandemic, the HIDOE is on schedule to secure a contract by July 15, 2020. Funding sources have been identified and the HIDOE intends to apply for a Competitive Grant for State Assessments in order to expand professional development and collaboration opportunities for participating teachers and relevant stakeholders. A mandatory training for participating teachers will occur near the opening of the school year and teachers may begin using the CBAS soon thereafter.</p> <p>The use of the CBAS results for the 2020-2021 school year is dependent on whether schools are operating normally next year. HIDOE hopes that it will begin incorporating CBAS results on Family Reports as soon as verified results as possible.</p>																		
<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</p>	<ul style="list-style-type: none"> Evidence requested in section (b)(7) above is also needed to satisfy this requirement. 	<p>[Please see b(2)(i) response above]</p> <p>Aggregated numbers for participating schools (below) show that subgroup participation is similar to demographic makeup of the whole school population, thus the participating sample does adequately include members of indicated subgroups</p> <table border="1" data-bbox="1115 1195 1866 1279"> <thead> <tr> <th>Year</th> <th>Subgroup</th> <th>Enrollment</th> <th>Free/Reduced Lunch</th> <th>IDEA</th> <th>ELL</th> </tr> </thead> <tbody> <tr> <td>2019-20</td> <td>Gr. 4 & 8 Aggregate</td> <td>3064</td> <td>41.84%</td> <td>10.97%</td> <td>9.60%</td> </tr> <tr> <td>2019-20</td> <td>School Aggregate</td> <td>13776</td> <td>43.83%</td> <td>11.09%</td> <td>12.30%</td> </tr> </tbody> </table> <p>The overall scale scores of non-participating school and pilot schools (using the shortened CAT) are comparable and can be used together for purposes like school classification as “in need of targeted support”,</p>	Year	Subgroup	Enrollment	Free/Reduced Lunch	IDEA	ELL	2019-20	Gr. 4 & 8 Aggregate	3064	41.84%	10.97%	9.60%	2019-20	School Aggregate	13776	43.83%	11.09%	12.30%
Year	Subgroup	Enrollment	Free/Reduced Lunch	IDEA	ELL															
2019-20	Gr. 4 & 8 Aggregate	3064	41.84%	10.97%	9.60%															
2019-20	School Aggregate	13776	43.83%	11.09%	12.30%															

relative to non-participating schools so that the SEA may validly and reliably aggregate data from the system for purposes of meeting requirements for--

(i) Accountability under sections 1003 and 1111(c) and (d) of the Act, including how the SEA will identify participating and non-participating 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.

etc.

Tables similar to the ones below will be generated for participating and non-participating classes for comparisons to ensure comparability. In addition, data will also be disaggregated by subgroups.

Longitudinal aggregated Smarter Balanced scores for all schools in the Longitudinal Data System (LDS):

SBA Language Arts Overall Proficiency ▾

Language Arts	2015-2016	2016-2017	2017-2018	2018-2019
Exceeded Standard	20.43%	19.87%	22.93%	23.27%
Met Standard	30.42%	30.01%	30.65%	30.36%
Nearly Met Standard	23.45%	23.01%	22.03%	22.07%
Not Met Standard	25.70%	27.11%	24.39%	24.30%

SBA Math Overall Proficiency ▾

Math	2015-2016	2016-2017	2017-2018	2018-2019
Exceeded Standard	18.00%	18.70%	19.50%	20.23%
Met Standard	23.59%	23.00%	22.57%	22.30%
Nearly Met Standard	28.71%	27.78%	27.39%	26.27%
Not Met Standard	29.71%	30.51%	30.54%	31.20%

Application Selection Criteria	Required information from the SEA	HIDOE Response
<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</p>	<ul style="list-style-type: none"> Evidence that the State has a plan that includes 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 initially participating schools as a baseline (e.g., how is school or regional leadership engaged in identifying schools to participate in the innovative pilot assessment). 	<p>HIDOE will conduct an annual internal evaluation of the project. Results of that evaluation will be used to inform project activities for upcoming years. Knowing that Grade 4 and Grade 8 were scheduled for NAEP testing in SY 2020-21, Grade 4 ELA and Grade 8 mathematics was selected as a starting point for the HICAP. The NAEP results for SY 2020-21 may serve as another comparison data set to gauge quality of initial implementation, professional development activities, and teacher designed tasks.</p> <p>In addition to the incentive to administer a shortened, summative CAT in place of the full version of the State’s assessment, participating teachers will be provided with a classroom set of laptop computers and a reliable Internet connection so that each student may be administered computer-based, teacher-created assessments throughout the school year. Teachers are encouraged to have students use the laptops during daily instruction and to explore differentiation strategies for various types of learners. It is anticipated that distance learning will continue for some time and the laptops may be used to administer CBAS assessments remotely. This is an important feature that will incentivize participation in the pilot program.</p>

<p>based on enrollment of subgroups of students described in section 1111(c)(2) of the Act and student achievement.</p> <p>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 initially participating schools as a baseline.</p>		
<p>(b)(2) The extent and depth of SEA, including each SEA in a consortium, and LEA capacity to implement the innovative assessment system considering the availability of technological infrastructure; State and local laws; 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</p>	<ul style="list-style-type: none"> • Evidence of the strategies HIDOE is using, or will use, to mitigate risks and support successful implementation of the local assessment component of the innovative assessment. 	<p>Professional development will be provided for participating teachers and support staff on how to build, administer, and score high-quality classroom-based assessments. HIDOE’s Assessment Section’s staff provides documentation (written and online training sessions/webinars) and annual face-to-face training sessions for test coordinators and members of school assessment teams.</p> <p>The Center for Assessment will support HIDOE's efforts in creating high-quality professional development opportunities to enhance teachers’ assessment literacy and capacity for professional practice.</p> <p>The HIDOE, in collaboration with the Center for Assessment, envisions multiple training sessions where teachers need to develop a classroom-based assessment (task), administer it, analyze student work, and subject the assessment to peer and expert evaluation.</p>

<p>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>The various test vendors also provide customer support (phone and email) to respond to questions from the field regarding access and technical support.</p> <p>The Assessment Section also produces a weekly newsletter covering all aspects of statewide testing including test windows, instructions with links, and technological updates that are sent to test coordinators, technology coordinators, and school administrators, as well as other state and district personnel involved with testing.</p> <p>In addition, the Assessment Section performs quality assurance and assessment monitoring site visits (see Appendix Q) to ensure school compliance with procedures and practices outlined in the state test administration manual.</p> <p>Each test vendor is responsible for maintaining a test delivery system for delivering assessments to students in a secure manner (e.g., through a secure browser) and for online test setup and monitoring by test administrators.</p> <p>Test vendors are also responsible for maintaining systems that, in the event of power or internet failure, capture student answers and store them for upload when connectivity is restored. Test vendors also ensure that their test delivery systems allow for the provision of accommodations such as text-to-speech, large print and other accessibility features as appropriate for students. For students who are unable to access the online platform, a system of test delivery in a paper format is made available. These systems must be compliant with the Family Educational Rights and Privacy Act (FERPA) and applicable HIDOE student privacy laws and guidelines.</p>
--	--	---

<p>(d)(2) The strategies the SEA or consortium has developed and will use to familiarize students and parents with the innovative assessment system;</p>	<ul style="list-style-type: none"> ● Evidence that there are plans to make various materials accessible to all parents, specifically for: <ul style="list-style-type: none"> ○ Those parents without Internet access. ○ Parents who have limited English proficiency. ○ Parents with a disability as defined by the Americans with Disabilities Act (ADA). 	<p>The HIDOE will host a virtual meeting for all HICAP pilot program participants including parents, students, teachers, principals and other stakeholders. Accessibility features will be provided for participants who are in need of support.</p> <p>The HIDOE provides parents/guardians with a hardcopy report (“Family Report”) of student performance on the statewide summative assessments. The HICAP family report will maintain the format and detail that has been provided on HIDOE’s reports since the first administration of statewide computer adaptive tests in SY 2010-11.</p> <p>Informational brochures, parent letters, and Parent Information Booklets will be produced in the 14 most common languages spoken in the state.</p>
<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</p>	<ul style="list-style-type: none"> ● Evidence of a detailed description of 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) HIDOE has developed, or plans to develop, in order to validly and reliably score local assessment items. 	<p>The standardized shortened summative CAT will be administered using the Smarter Balanced pool of items. The assessments will be machine-scored using the Smarter Balanced specifications for automated scoring.</p> <p>The authentic assessments that are created and administered by teachers using the Classroom-Based Assessment System are not required to meet the same level of technical quality as those that are used to generate the overall score that is used for accountability purposes. However, HIDOE will employ similar comparability procedures that are used with New Hampshire’s Performance Assessment of Competency Education IADA to ensure both within district interrater reliability and cross-school/district comparability, The CBAS will use machine scoring technology to provide results immediately upon completion of an assessment that includes items from a bank that are calibrated for</p>

<p>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>machine scoring, including full-write and short answer items. Teacher-created items that require hand-scoring will be scored by the teacher or a group of teachers that are collaborating to develop a common assessment task.</p>
<p>(e)(1) The strength of the proposed evaluation of the innovative assessment system included in the application, including whether the evaluation will be conducted by an independent, experienced third party, and the likelihood that the evaluation will sufficiently determine the system’s validity, reliability, and comparability to the statewide assessment system consistent with the requirements of 34 CFR part200.105(b)(4) and (9);</p>	<ul style="list-style-type: none"> ● Evidence that the proposed third-party evaluation will address the innovative assessment system’s validity and reliability, specifically plans to independently verify alignment of the two assessments (the short summative State test and the local assessments) with the State’s academic content standards. ● Evidence of plans to address the comparability and alignment between the short Statewide summative assessment and the local assessments. 	<p>The Shortened Summative Assessment (SSA) is based on the Smarter Balanced blueprint, which has been proven to produce scores comparable to the full Smarter Balanced using only the CAT portion of the assessment. The CBAS does not need to be included as the SSA will yield valid and reliable scores comparable to the full Smarter Balanced assessment composed of both the CAT and performance task (PT). Thus, no third-party evaluation of technical validity or interrater reliability is necessary as the CBAS will not be used as part of accountability reporting, at least initially until results can be verified as meeting key technical quality criteria.</p> <p>The CBAS will be used in place of the PT to provide teachers a means to produce vetted, shared classroom assessments that will allow them to access the deepest levels of knowledge, as well as measure, identify, and address gaps in learning identified by discrepancies between the SSA, CBAS scores, and grades.</p>

DAVID Y. IGE
GOVERNOR



DR. CHRISTINA M. KISHIMOTO
SUPERINTENDENT

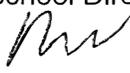
STATE OF HAWAII
DEPARTMENT OF EDUCATION

P.O. BOX 2360
HONOLULU, HAWAII 96804

OFFICE OF STRATEGY, INNOVATION AND PERFORMANCE

September 18, 2019

TO: Deputy Superintendent
Complex Area Superintendents
Public Charter School Executive Director
Principals (All)
Public Charter School Directors (All)

FROM: Rodney Luke 
Assistant Superintendent

SUBJECT: Quality Assurance and Assessment Monitoring Site Visits

The Assessment Section conducts annual Quality Assurance (QA) site visits to meet with school level staff associated with test coordination and administration to identify and discuss best practices and areas in need of support in the coordination and administration of statewide assessments. These visits are informational in nature, and provide opportunities for Assessment Section staff and designees to assess the effectiveness of trainings and to obtain feedback on how support may be improved in the future. Feedback and shared best practices are welcome, as it allows Assessment Section staff members to share these in trainings so that other schools may benefit from these practices. QA site visits are also a great opportunity for school staff to learn about statewide testing requirements and to effectively plan for implementation.

Schools are selected for QA site visits based upon several factors. Some schools are selected randomly while others are selected based upon test administration practices noted in prior years. An Assessment Section staff member works with the school test coordinator to plan the QA site visit that may occur at any time during the school year. A summary of the QA site visit will be made available to the principal upon request.

The Assessment Section also conducts annual assessment monitoring site visits during statewide assessment testing windows to observe test administrations and to gather information from school test coordinators about state assessment administration training, management and practices. These visits provide an opportunity for Assessment Section staff and designees to see the testing process in action as well as to identify practices and policies where improvements can be made. The information gathered during these visits is used to improve the testing experience of all those involved while at the same time ensuring the validity and reliability of the assessments.

During assessment monitoring site visits, Assessment Section staff and designees observe the extent to which school practices align with state assessment policies and guidance. For example, monitors will note whether:

- All cell phones and other electronic devices are stored in a secure location and are not accessed at any time by students or adults during testing;
- Bulletin boards, posters and other materials containing content-related information are not visible in the testing room;
- Students are properly seated in the testing room and engaged in the testing process;
- Test administrator follows guidance provided in the Test Administration Manual (TAM) (e.g., reading the appropriate test directions);
- Documentation of test administrator/proctor trainings such as sign-in sheet lists indicating dates and time of the trainings;
- Students who require testing accommodations are provided those accommodations;
- Students who require designated supports (or other accessibility features, e.g., administrative considerations for the WIDA ACCESS) are provided those supports;
- Test administrators and proctors are not interfering with or prompting student responses in any manner;
- Test administrators and proctors are not restricting student progress during testing sessions (i.e., students should not be told to answer only a limited number of questions within a testing session);
- Test administrator and proctor actively engage in monitoring of students during testing; and
- Proper distribution, collection and storage of secure test materials, such as test tickets and scratch paper.

Schools are selected for assessment monitoring site visits based upon several factors. Some schools are selected randomly while others are selected based upon test administration concerns noted in prior years. Since any school may be selected for a site visit, all public and public charter schools should anticipate the presence of an assessment monitor during the administration of the Smarter Balanced, HSA-Alt, HSA Science (NGSS), Biology 1 (NGSS) EOC Exam, KĀ'EO, The ACT, NAEP and the ACCESS for ELLs assessments at any time during the respective testing windows. Assessment monitoring site visits may be unannounced or arranged with minimal lead time to ensure that observations are made of typical practice. Observations of typical test practice will be used to improve test coordinator and test administrator trainings. A summary of the site visit will be made available to the principal upon request.

If you have any questions, please contact Bruce Hirotsu, Assessment Section, at (808) 307-3636 or via bruce.hirotsu@k12.hi.us.

RL:bh

c: State Public Charter School Commission
Assessment and Accountability Branch

Appendix T: Performance Assessment Training

Performance Assessment Training

The professional capacity building plan is conceptualized as an initial half-day content-specific webinar with the cadre of experts each year to ensure common understandings of each discipline and provide an opportunity to brainstorm direction for task designs using the Hawai'i content standards and Smarter Balanced claims/assessment targets.

Most of the professional capacity building will occur during a series of four Performance Assessment Institutes that will take place from October to mid-Summer each year. These Performance Assessment Institutes are anticipated to take place over two days per Institute. Teachers will be paid a stipend to participate. The development of high-quality common performance tasks is grounded in the training of the cadre of experts and groups of teachers. The HIDOE and Center for Assessment will lead the task development work with input from the cadre of experts. This model allows the HIDOE and cadre of experts to be poised to continue the work after the planning grant ends. Figure 1 below shows the proposed trajectory and content of the four Performance Assessment Institutes.

Figure 1

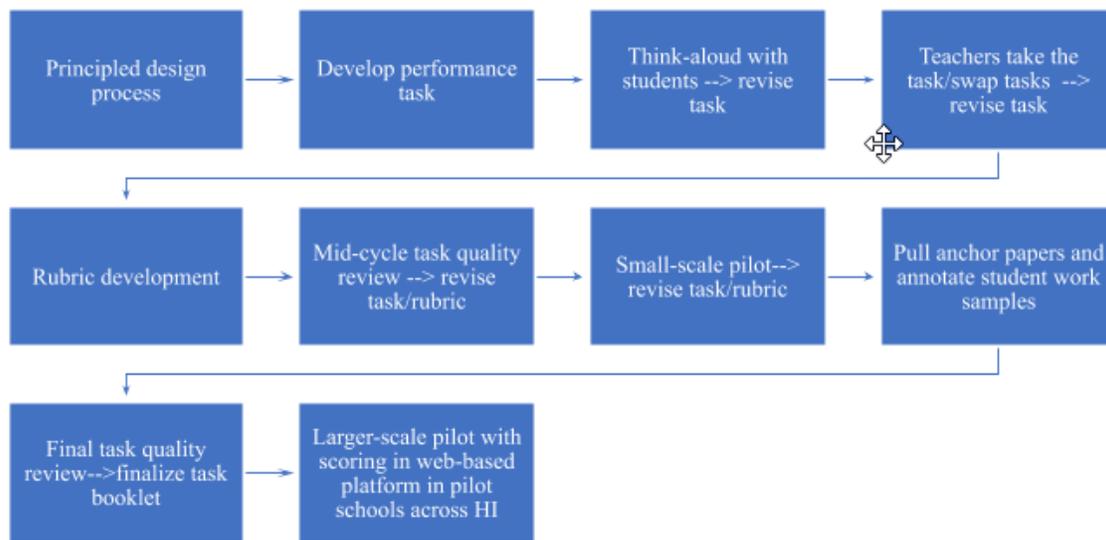
Proposed Trajectory and Content of the Performance Assessment Institutes



The Performance Assessment Institutes will lead participants through a principled assessment design process (i.e., evidence-centered design) that incorporates the principles of Universal Design for Learning (UDL). Figure 2 illustrates the common performance task development and pilot-testing process described in detail after the figure.

Figure 2

Common Performance Task Development and Pilot-Testing Process



Principled assessment design means the task is developed based on (a) what students should know and at what depth of knowledge; (b) what evidence is necessary to demonstrate that the student has the desired knowledge; and (c) what tasks will allow students to demonstrate and communicate the desired knowledge.

A “backward design” model (Wiggins & McTighe, 2005) performance task template will be created to provide guidance on the characteristics of a high-quality task and expectations. This template is used by educators to initially **develop multiple performance tasks**—some of which may become instructional tasks—for each grade and subject area, which are designed to provide data on how students are progressing toward the state content standards.

In line with principles of UDL, common task developers will consider during the design phase the extent to which the performance task provides students with (a) *multiple means of representation* to give learners various ways of demonstrating information and knowledge gained through instruction, (b) *multiple means of expression* to provide learners alternatives for

demonstrating what they know, and (c) *multiple means of engagement* to tap into learners' interests, challenge them appropriately, and motivate them to learn.

Task Quality Review, Integration with Web-Based Classroom Assessment Platform, and Larger Scale Piloting

Once the common performance tasks are initially developed, cognitive laboratories (also known as **think aloud protocols**) will be used with students to collect evidence about task quality and the thinking processes that students employ when interacting with the task. Tasks are then revised based upon student feedback.

Teachers will then take the performance task themselves and swap performance tasks in order to examine task quality and gather suggestions for revision. **Rubrics will be developed** to describe student performance relative to Hawai'i's content standards.

The HODOE, with support from the Center for Assessment, will then model for the cadre of experts how to conduct a **mid-cycle task quality review** of the tasks and rubrics using a Task Review Tool. This tool will be developed using the criteria for high-quality assessments from the *Standards for Educational Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). These criteria include considerations around alignment, cognitive complexity, fairness, bias and cultural sensitivity, use of appropriate text/visual resources, and quality of scoring guidelines and criteria.

Particular attention will be paid in the review to the appropriate specifications around accommodations for students with disabilities and English language learners using the principles of UDL. Specifically, common performance tasks will be reviewed based on whether they measure student skills that are outside the intended construct, use extraneous words that potentially distract students from the main learning target of the task, use idioms, or culturally-

specific language, crowd text and/or graphics too closely on the page, and/or use graphics that require certain levels of visual acuity to understand. This Task Review Tool will identify areas of strength and provide recommendations for revisions. This feedback will then be used by the cadre of experts to revise the tasks as necessary prior to small-scale pilots.

Teachers will conduct **small-scale pilots** to evaluate and refine task quality. The groups of teachers involved in common task development will suggest revisions to the tasks and/or rubrics. Revisions will be made to the tasks and/or rubrics as necessary by the cadre of experts. The revised tasks will then be re-piloted in some classrooms and **anchor papers identified and annotated** to support reliable scoring.

At the end of the common performance task development process, the HIDOE, with support from the Center for Assessment, will provide a **final review of all task materials** (i.e., task booklet) and approve the task for operational use—or send the task back to the cadre of experts for revision until the task is approved.

The set of common performance tasks created in each grade/subject area will be integrated into the web-based classroom assessment platform and create a bank of common performance tasks aligned to Hawai'i's content standards that can be used to support the classroom assessment component of HICAP and local assessment purposes throughout the year. The final task materials will then be used by the pilot schools to conduct a **larger-scale pilot** with pilot scoring within the web-based platform/standards-based grading system in the spring. This will allow the HIDOE to resolve any issues with the web-based platform/standards-based grading system and collect data that could be used to determine how best to report the common performance tasks alongside the shortened summative component of HICAP. This cycle repeats each year with new cohorts of cadres of experts and teachers.

