



Hawai'i Department of Education

Appropriate Use of SY 2020-21 HSAP RESULTS

The Hawai'i Statewide Assessment Program's (HSAP) portfolio of assessments includes the end-of-year Smarter Balanced, HSA Science, KĀ'EO, and HSA-Alt assessments. The Hawai'i Department of Education (HIDOE) encourages educators to use the HSAP test results for appropriate purposes that recognizes the impact of the COVID-19 pandemic. This document provides background information about the HSAP, offers general guidance about interpreting and using the assessment results this year, and concludes with vignettes of various case studies to illustrate key aspects of the guidance.

BACKGROUND

Use end-of-year test results to support teaching and learning as part of HIDOE's COVID-19 recovery plan

It is important to keep the 2020-21 HSAP test results in context. There were different teaching and learning strategies employed and every effort was made to maximize learning.

- Teachers, school leaders, and parents have gone to great lengths to help students learn as much as possible last year despite unprecedented challenges.
- End-of-year test scores provide important information and should be interpreted thoughtfully, and always in combination with other information, such as students' grades, teachers' observations, and attendance patterns.

The HSAP end-of-year assessments provide useful information about what students may have learned this past year

The HSAP summative assessments provide useful information about student learning for education leaders during a year when teaching and learning looked very different.

- A crucial goal for HIDOE is to understand what students learned during a given year, even when learning looked different from previous years. End-of-year assessments provide information about student learning to state, complex area, school, and other education leaders to help them improve equity and prioritize resources where they might be needed most.
- End-of-year testing is one piece of the HIDOE assessment system (including interim assessments, [Tools for Teachers](#), [Interim Connections Playlists](#), etc.). The full systems also include tools and resources for educators specifically designed to be easy to find, to save time supporting student development, and to provide actionable information to advance student learning.
- In order to build and strengthen an effective, equitable education system, educators need state tests that provide comparable information across complex areas, complexes, and schools.
- HIDOE is committed to having accurate information about student progress, each and every year. We had to make tough decisions this year about how and when to test students to minimize interruptions to learning, while still yielding important information.
- HIDOE collaborated with other states and assessment experts to ensure that we used the best possible assessment practices during a year that has not been normal for anyone, especially teachers and students.



Hawai'i Department of Education

- To this end, HIDOE offered shortened Smarter Balanced Assessments for use in school year (SY) 2020-21, allowing the flexibility to collect high-quality end-of-year information that still produced useful information for educators, leaders, and policymakers.

GUIDANCE

Considerations/Cautions

The primary considerations associated with interpretation of the SY 2020-21 HSAP results can be broadly grouped in to three categories:

- ***Diminished opportunities-to-learn (OTL).*** Opportunity-to-learn refers the conditions and resources provided to schools to enable students to succeed. Students experienced disruptions in OTL this year compared to “typical” years.

Students experienced various learning disruptions this school year, which may have included first-time experiences with remote learning, limited access to internet and technology to allow full participation in remote learning, reduced instructional time, and varying degrees of supports and resources when learning at home. Some students, such as those from low-income families or English learners, were likely hardest hit by the COVID-19 pandemic due to having fewer resources and supports.

OTL indicators should address more than interactions between students and teachers—as important as those are—but on larger systematic issues such as high-quality internet connectivity, availability of devices, and other critical resources, which were all brought into sharp relief during COVID. Sharing OTL information provides an important perspective for anyone reviewing test results and making comparisons with prior year results.

- ***Changes to the test design or administration.*** HIDOE relied on a streamlined test blueprint for the Smarter Balanced Assessments to reduce test-taking time, but this difference must be considered when comparing to previous years.

HIDOE made additional changes to tests this school year to accommodate test taking from home, shifting school schedules, and evolving health and safety requirements. Modifying testing modalities and the requirements around test administration complicates comparisons and raises concerns that results from tests administered in-person cannot be meaningfully compared with those administered remotely.

- ***Changes in tested population.*** This is one of the most critical things to understand when interpreting test scores this year. Even small changes (e.g., 5%) in the population and specific subgroups of students tested in 2021 compared to 2019 can lead to misunderstanding of the state test results.

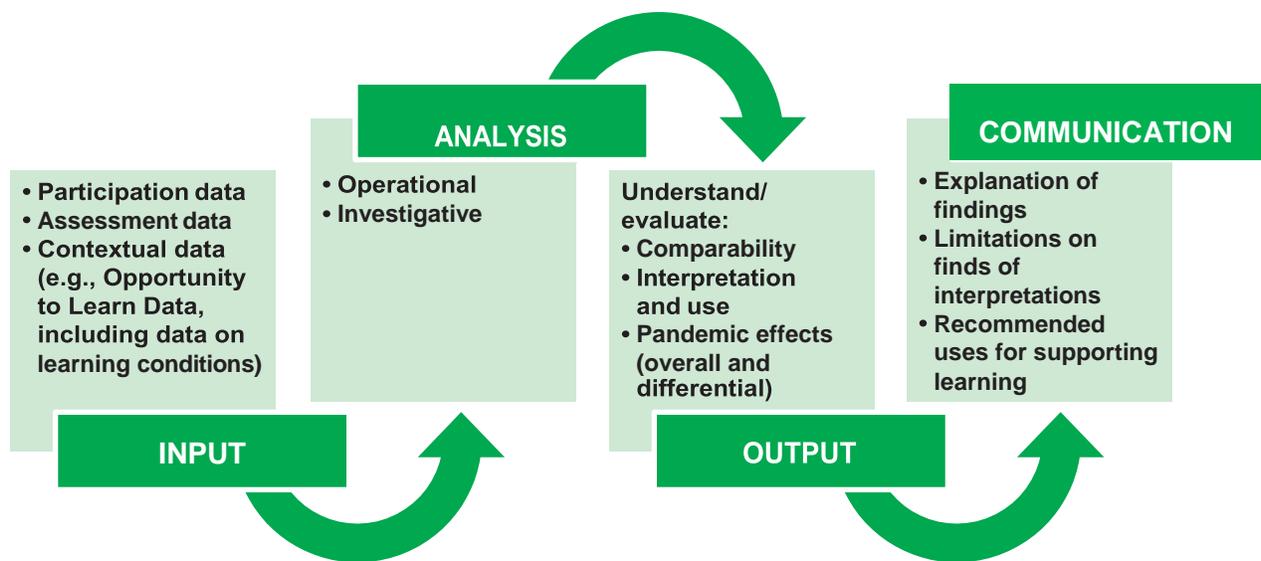
Some students were able to take tests this school year while others were not due to safety concerns, challenges with technology, or other barriers. This means that some participation rates for districts, schools, or student groups are lower than in past years. As participation rates decrease, challenges with interpreting results increase. In addition, the wide availability of different learning settings—in-person, remote learning, or hybrid—means that students had varying access to take state tests. Thus, some student groups will be overrepresented in the results and others may be underrepresented.



Hawai'i Department of Education

Federal assessment requirements call for a 95% participation rate of students eligible to test. The purpose of this requirement is to ensure test results are representative of all students who should be testing. When participation rates drop, concerns over which students are not testing arise. Alternatively stated, when participation rates drop, are test results representative of all students? Are only certain types of students testing? Are certain race/ethnicities testing at a lower rate than others? Are certain disadvantaged groups such as economically disadvantaged, English Learners, and Students with Disabilities, testing at lower rates than non-disadvantaged students? Participation rates give us an indication that possible disproportionate representation comprises test results. At minimum, participation within student subgroups should be examined individually and taken into consideration when reporting or making inferences of those results.

Changes in OTL, test design and administration, and the number and percent of students participating in a test must be considered when analyzing the 2020-21 HSAP test results. Therefore, HIDOE and its assessment partners are prioritizing analyses that will most likely inform understanding of the effects on learning. The illustration below provides the logical flow of analyses based on SY 2020-21 results that begins with inputs and proceeds through to the communication of results.



Analysis Framework Logic Model (Dadey et al., 2021)

Remember, a single test score does not provide a complete or precise measure of student achievement. Educators and school leaders must consider other measures of student achievement, especially those that are closer to the classroom.

Also, consider how the conditions for learning, which may have been disrupted by the pandemic and may influence performance. Caution is urged in interpreting summary results when participation rates are low. The standard participation threshold is 95%. When participation falls below that level, inferences about overall performance are uncertain, especially when the composition of test takers is different compared to the underlying school population in 2021 and/or compared to the test population in prior years.

Local school and network leaders should consider the following guidance on participation rates when presenting and using their assessment results.



Hawai'i Department of Education

Participation

High 95%+	When both overall and subgroup participation rates are high, results can be used to inform decision-making. Even with high participation rates, educators and leaders still need to consider the extent to which the standard curriculum was covered and the extent to which students were engaged in the learning. The shortened Smarter Balanced Assessments allow for comparison of overall proficiency, however, the claim level indicators are not reliable for the shortened blueprint and should not be used for decision-making.
Medium 85-94%	When overall and subgroup participation is above 85% but below the 95% target, results can still be used to inform decision-making, however, additional questions need to be addressed. For example, how representative is the group of students who tested compared to the population as a whole? Is the group that tested generally high or low achieving? Do they differ in terms of demographics (SPED, EL, disadvantaged, etc.) from the population? It is important to review multiple sources of data such as formative assessment results, attendance, and OTL.
Low 75-84%	Educators and leaders should be cautious when analyzing results of a low participating group. If the size of the population is relatively small (less than 50) and the participation rate is low then it is imperative to consider how representative the sample of students tested is of the whole (population). The ratios for subgroups need to match those of the population otherwise results are not comparable to previous years. The degree to which content standards were covered and students were engaged must also be taken into consideration.
Insufficient <75%	Participation rates of less than 75% generally are not sufficient to make comparisons with previous years' results without relying on additional statistical techniques. Results in the aggregate may not be representative of the population and decisions should be made very cautiously. Individual student results may be used to inform understanding of individual student achievement (this is true regardless of participation rates), however, questions with regards to content coverage and engaged learning must be addressed.

Case Scenarios

I. Large Population, High Participation; High Content Coverage; Mostly In-Person Learning

Population (count of students)	SBA Participation (percent of students)	Curriculum Coverage (Hawai'i Common Core)	Distance Learning (percent of year)
200	95% overall and for all subgroups	90%	40%

In this case, the large population and high rate of participation allows for comparison with previous administrations at the overall score and target levels. The shortened Smarter Balanced Assessment generates reliable overall and target scores. *However, the claim level scores are not comparable to previous years due to fewer test questions for each claim.* The "At/Near" level for each claim will most likely be larger this year due to the greater standard error of measurement; claim level indicators are still meaningful but not comparable to previous test administrations.

II. Medium Population, Medium Participation, Medium Content Coverage, Mostly Distance Learning

Population (count of students)	SBA Participation (percent of students)	Curriculum Coverage (Hawai'i Common Core)	Distance Learning (percent of year)
100	85% overall and for all subgroups	80%	90%

The overall participation rate for this population – including subgroups – is borderline for comparison with previous administrations and caution is advised if data are used for decision-making purposes. The reduced content coverage and high rate of distance learning must also be taken into consideration which will make it difficult to accurately determine cause and effect. It is important to consider the demographics of the students tested and compare them with the population as a whole before making any comparisons to previous year's overall proficiencies.



Hawai'i Department of Education

III. Small Population, Low Participation, Low Content Coverage, Mostly Distance Learning

Population (count of students)	SBA Participation (percent of students)	Curriculum Coverage (Hawai'i Common Core)	Distance Learning (percent of year)
50	75% overall and for all subgroups	60%	80%

The combination of a small population and low participation makes it inappropriate to make any comparisons with previous years' results. Only individual student results may be used to inform understanding of the extent to which an individual student met expectations for learning. Consideration needs to be given to the extent to which the student was taught the breadth and depth of the content standards and the extent to which the student was engaged in the learning process – particularly given the large percentage of time spent distance learning.

Example

Aloha Middle School's overall proficiencies for the Smarter Balanced ELA Assessments are provided in the table below for the past four school years that the SBAs were administered (NOTE: the SBA was not administered in SY 2019-20). There were approximately 300 students who took the ELA SBA over the past four administrations and the subgroups population was approximately 50 students each of the four years of test administration.

ELA Smarter Balanced Assessment Proficiency				
School Year	Proficiency Rate (Schoolwide)	Participation Rates (Schoolwide)	Proficiency Rate – High Needs Subgroups	Participation Rates – High Needs Subgroups
2016-17	45%	99%	33%	96%
2017-18	52%	97%	35%	97%
2018-19	54%	98%	37%	96%
2020-21	60%	90%	20%	50%

At first glance, the schoolwide overall percent proficiency of 60% and a high rate of participation of 90% would suggest that students were successful last year in the mostly distance learning model at Aloha Middle School - even though the high needs subgroups have a large gap that needs to be addressed. However, upon further inspection we see that the participation rates of the high needs subgroups were 96-97% in the previous three administrations. The low participation rate of 50% of the high needs subgroups in SY 2020-21 skews the schoolwide overall proficiency rate and provides a “false positive” of the overall performance of the students at the school. Additional context is needed in order to conduct meaningful analysis of the data at a more granular level (e.g., grade level, class, and student). An example of how to add further context is through taking a virtual data dive with the [Multiple Measures Student Screening \(MMSS\)](#) report in LDS. With this report you are able to analyze individual student risk factors for attendance, discipline, report card marks, and universal screeners in order to illuminate, confirm, or dispute what you learned from your initial analysis of the ELA SBA data.

Resources/References

Domaleski, C., D’Brot, J., Marion, S., & Boyer, M. (2021). [Sensible Reporting of Spring 2021 State Assessment Results](#). Dover, NH: The National Center for the Improvement of Educational Assessment.

Arnold, J. & Crane, E.W. (2021). [Making Decisions: Using Assessment Data in the Time of COVID-19 and Beyond](#). WestEd.

Dadey, N., Keng, L, Boyer, M., & Marion, S. (2021). [Making Sense of Spring 2021 Assessment Results](#). Dover, NH: The National Center for the Improvement of Educational Assessment.