2019 No. 102

California Assessment of Student Performance and Progress (CAASPP)

2019 Independent Evaluation Report

Prepared for:
California Department of Education
Assessment Development and Administration Division
1430 N Street, Suite 4401
Sacramento, CA 95814-5901

Prepared under:
CN180100

Authors:
Michele M. Hardoin
Rebecca Norman Dvorak
Arthur Thacker
Justin Paulsen
Monica Gribben
Kristina Handy

Editors:
Sheila Schultz
Lauress Wise
Christa Watters

Date:
October 31, 2019

Headquarters: 66 Canal Center Plaza, Suite 700, Alexandria, VA 22314 | Phone: 703.549.3611 | https://www.humrro.org

California Assessment of Student Performance and Progress (CAASPP)

2019 Independent Evaluation Report

Table of Contents

Table of Contents (cont.)

List of Appendices

List of Tables

Table of Contents (cont.)

List of Tables

List of Figures

Executive Summary

Pursuant to California *Education Code* (*EC*) Section 60649, the Human Resources Research Organization (HumRRO) is continuing its independent evaluation of the California Assessment of Student Performance and Progress (CAASPP) System. The scope of the current evaluation is to conduct three research studies from July 2018 through December 2020 and provide objective technical advice and consultation on activities related to the implementation of specific components of the CAASPP.

The 2018–20 CAASPP Evaluation Plan is presented in HumRRO's publicly available *2018 CAASPP Evaluation Report* (https://www.cde.ca.gov/ta/tg/ca/documents/caaspp18evalrpt.pdf). The report consists of the CAASPP System's theory of action (CDE, 2018a) and detailed plans for each evaluation study. The plan also includes a timeline for major study milestones; the timeline is based on CDE priorities and the anticipated dates of operational administration of assessments.

This annual report covers the activities HumRRO conducted during the 2018–19 school year for each of the following studies:

- Impact on Instruction and Student Learning Case Study (hereafter, Impact Case Study)
- California Science Test (CAST) Alignment Study
- California Alternate Assessment (CAA) for Science Alignment Study

The multiple systems that form the California assessment environment are complex. Across the state, local educational agencies (LEAs), schools, and teachers continue to implement the various components of the expanded CAASPP System, which now includes two new science tests (the CAST became operational in 2018–19, and the CAA for Science will become operational in 2019–20); the new English Language Proficiency Assessments for California (operational as of 2017–18); and the California Spanish Assessment (operational as of 2018–19). The CDE, its testing contractors, and the Smarter Balanced Assessment Consortium continue to enhance the Smarter Balanced components.

The CAASPP System includes assessments as well as resources to help teachers, administrators, students, and parents prepare for the assessments and understand subsequent results. HumRRO's Impact Case Study addresses three well-established Smarter Balanced components: Summative and Interim Assessments for English language arts/literacy (ELA) and mathematics, and the Digital Library (DL). The two alignment studies address the newest CAASPP components, the CAST and the CAA for Science. This executive summary provides an overview of each study, detailing progress made to date in terms of data collection, data analysis, and summarization of findings and recommendations. HumRRO will complete year two of the Impact Case Study, both alignment studies, and reports on each study during the remainder of the contract.

Impact Case Study

Overview

According to the CAASPP program theory of action, the Smarter Balanced components—working together to accurately assess student achievement relative to grade-level curriculum standards, the Common Core State Standards (CCSS)—provide information to educators to improve instruction and thus improve student achievement. The first year of the Impact Case Study examined LEAs that are implementing Smarter Balanced Interim Assessment Blocks (IABs) for ELA or mathematics (in addition to the mandated summative assessments) to explore how the theory of action may be driving efforts to improve student achievement.

HumRRO employed a case-study approach during the first year of this two-year study. HumRRO defined a case as an LEA that had at least a modest threshold of use of the Smarter Balanced IABs in 2017–18 and planned to continue its use during 2018–19. During the 2018–19 school year, HumRRO collaborated with seven LEAs, including one direct-funded charter, encompassing 19 schools.

The primary goal of the Impact Case Study is to elicit concrete examples of how and why specific CAASPP components (i.e., Smarter Balanced components for ELA and mathematics) are used, their impact on instruction and student learning, and the perceived benefits, strengths, and challenges of using the components. For this first year of the study, the evidence we collected about "impact" was related mostly to policies and practices for implementing CAASPP components. The very specific group of LEAs, schools, and educators that participated in our study provided very few examples of impacts on student learning, and the impacts on instruction we identify are not generalizable beyond our small sample. However, we feel the information from this first year is meaningful for the CDE and for LEAs as they consider how CAASPP components can be used in combination with other resources and what aspects might need to be improved. Chapter 2 presents the 13 research questions addressed by this study; descriptions of the LEA sample selection, data collection activities, and data analysis methods; and overall findings across LEAs, by research question. Appendices present in-depth and summary findings, by LEA.

Summary of Findings

This section provides a high-level summary of the findings (across the sample of LEAs and schools in the study) associated with the use of three well-established Smarter Balanced components: summative assessments; interim assessments (IAs), which include shorter Interim Assessment Blocks (IABs) and longer Interim Comprehensive Assessments (ICAs), and the DL.

Summative Assessments

School staff participating in the study reviewed summative assessment data from the prior year (2017–18) during the first semester of the 2018–19 school year. Some schools reviewed data as a school-wide team early in the year, while other schools did not do so until November or December. Delays in review of data were due to decisions made at the district level or confusion about the allowable uses of preliminary results. However, when scores were made available to districts in the Online Reporting System (ORS) in June 2018, the Deputy Superintendent of Public Instruction's letter to LEAs specifically stated the ORS results were not embargoed and encouraged use of the early results to inform educational programs and support local planning around the improvement of teaching and learning. The degree to which data were reviewed and used varied among schools. Almost all school leaders and teachers at the elementary and middle schools reviewed grade-level results of the percentage of students who fell into each overall achievement level for ELA and mathematics. Many also reviewed results by claim, and a few accessed target reports. Some teachers in our study had trouble recalling anything about the prior year's summative assessment scores and thus did not describe how the results influenced instructional activities. In contrast, some schools described how summative assessment scores were a central piece of evidence for identifying annual achievement goals, and in some cases the summative assessment scores influenced instructional foci and/or the selection of IABs to administer during 2018–19.

Interim Assessments

IAs were used by all schools in the study except one high school. Some LEAs mandated IA use, either by indicating the minimum number of IABs and/or ICAs to be administered per subject and grade level, or by mandating the specific IABs to administer. Other LEAs allowed schools and/or individual teachers or teacher groups to make these decisions. In LEAs with mandates, teachers were allowed to administer additional IAs.

Many teachers felt IAs were beneficial for preparing their students for the content, rigor, item types, and technology they would face on the summative assessments. Teachers believed exposure via IAs would benefit students' ability to demonstrate their knowledge and skills on the summative assessments. Some teachers saw additional benefits of IAs, finding them useful to identify gaps in student understanding and determine what content needed to be retaught. Some teachers chose to give the same IAB twice, as pre-test and post-test, to measure growth in student knowledge, though the CDE advises LEAs to be cautious in interpreting the results when IAs are used in this manner. Other teachers gave IAs only because of LEA-level mandates. There were mixed feelings on the utility of the IA Reporting System. In some cases, especially when results were accessed through the LEA's student information system, it seems teachers were not aware of the various features (e.g., reporting levels, item analysis, etc.) available to them through the IA Reporting System.

Digital Library

The study schools reported extremely limited use of DL resources. Most teachers were aware of the resources and had logged directly into the DL at least once; however, teachers noted time constraints, accessibility of sufficient materials through their curriculum or other familiar sources, and difficulty navigating through the system as reasons they were not using the DL. Only two teachers across the entire study indicated the DL was beneficial for classroom instruction in 2018–19. HumRRO did not investigate teachers' use of the alternate route to the DL resources via the IA Reporting System.

Summary of Best Practices

This section provides a high-level summary of a sample of the best practices evidenced among the collaborating LEAs and schools in response to the Impact Case Study research questions. The research questions addressed use of the three Smarter Balanced components studied (i.e., summative assessments, IAs, and the DL). For this report, HumRRO defined a "best practice" as an approach used by participating LEAs, schools, or teachers that (a) aligns well with the intended purpose of and guidance for implementing components within the CAASPP System and (b) resulted in educators having a positive experience using the CAASPP System to inform their teaching. We believe these practices may benefit other schools or LEAs, though we acknowledge there are multiple ways to achieve the goals of the CAASPP System. Additionally, schools and LEAs need to balance approaches to meet their available resources.

Across the studied LEAs and schools, HumRRO identified the following sample of best practices used by participating LEAs for successful implementation of the Smarter Balanced components:

- Provide support and training at the school and LEA levels for using CAASPP resources. Teachers and staff who attended CAASPP professional development or reviewed resources available online increased their comfort level with the CAASPP components, including hand scoring of IABs and using and interpreting assessment results.
- Provide leadership guidance and encouragement for using CAASPP components while allowing grade-level or content-area professional learning communities (PLCs) flexibility regarding what interim assessments and DL resources to incorporate into their classrooms.
- Facilitate school-wide data discussions to ensure teachers know how to access and interpret results, and how these data can inform instructional practices.
- Provide time and resources to support collaboration among grade-level and/or content-area PLCs to plan instruction and use formative assessments effectively.

Recommendations and Planned CAASPP System Improvements

HumRRO reviewed the full scope of study findings based on the perspective of the participants—a small number of teachers within a small number of schools in a small number of LEAs—to develop suggestions for the CDE to consider as part of its continuous improvement of the CAASPP System.

Based on the first-year findings across the case study LEAs, we offer the following recommendations. Some recommendations are already being addressed by enhancements the CDE will implement during the 2019–20 school year. Where applicable, recommendations are followed by brief descriptions of important CAASPP System improvements that will respond to areas of need described by LEA and school staff or observed by HumRRO.

Recommendation 1: The CDE should continue providing regional training opportunities and updated online resources for LEA- and school-level staff. The in-person trainings and CAASPP.org and CDE website resources are critical to helping educators throughout the state (a) accurately interpret Smarter Balanced summative and interim assessment results, (b) implement existing and new Smarter Balanced components, and (c) learn about enhancements to existing components.

Planned CAASPP System Improvements:

- The CDE will host a statewide 2019 California Assessment Conference in October 2019. The three-day conference will offer a variety of sessions for classroom educators to explore the connection between assessments and classroom instruction and to explore ways of using assessment resources for improved teaching and learning.

- Beginning September 3, 2019, educators will use a single username and password (i.e., single sign on) to access the various CAASPP and ELPAC online systems, including the Test Administrator Interface, Interim Assessment Systems (Viewing System, Hand Scoring System, and Reporting System), Online Reporting System (ORS), DL, and Practice and Training Tests.

- The CDE will release a new interface to the DL, currently referred to as the DL 2.0. The updated DL will address many of the concerns with the current DL. It is expected to be easy to use, will include step-by-step directions, and will be accessible (WCAG 2.1AA compliant). The DL 2.0 is being purposefully developed to align with Smarter Balanced grade-level claims and targets and provides options and ideas for differentiation and student access of content. Instructional resources will be embedded with the formative assessment process strategies. In addition, the DL 2.0 is aligned with new Smarter Balanced quality control criteria. Finally, the DL 2.0 resources will be specifically tied to Connections Playlists, tools that link interim assessment

results with teaching resources in the Digital Library to help optimize student learning.

Recommendation 2: Regarding interim assessments, the CDE should work with the Smarter Balanced Assessment Consortium to provide an expanded pool of ELA and mathematics tests, including multiple versions of existing IABs, ICAs for grades nine and ten, and shorter interim assessments that examine student achievement at the target level. Teachers using the existing interim assessments find them of high quality and requested more options for tests for classroom use.

Planned CAASPP System Improvements:

- New Smarter Balanced ICAs in ELA and mathematics will be available for administration to students in grades nine and ten in 2019–20, with different cut points for each grade level.

- New Smarter Balanced Focused IABs will measure one to three targets compared to up to eight targets measured by the current IABs. These focused IABs will measure smaller bundles of content to (a) give teachers a better understanding of students' knowledge and academic performance and (b) provide teachers with precise next steps for instruction. In addition to the more than 100 IABs already available to teachers, approximately 40 focused IABs are slated for release in 2019–20, followed by approximately 90 more over the following two school years.

Recommendation 3: Regarding the hand scoring requirements of some interim assessments, the CDE should explore how to address concerns related to the challenges some LEAs and schools have finding time for training and hand scoring. Some teachers in our sample who participated in hand scoring found it an excellent professional development activity, and others found instructional value in reviewing scored responses. However, constraints on time and resources often caused schools to decide against giving IABs that involve hand scoring. Perhaps the CDE could include an option for scoring via artificial intelligence techniques (currently in progress by Educational Testing Service, ETS). At the local level, support could take the form of (a) increasing the number of in-person hand-scoring training opportunities, (b) expanding the number of participants in such training, (c) providing teacher release time to engage in hand-scoring activities, or (d) sharing examples of teachers enthusiastic about their experiences with hand scoring (e.g., the CAASPP in Action series).

Recommendation 4: The CDE should encourage LEA and school leaders to provide local training opportunities, including time and resources, to help teachers (a) accurately interpret Smarter Balanced summative and interim assessment results, (b) implement existing and new Smarter Balanced components, and (c) learn about enhancements to Smarter Balanced components. LEA and school leadership receive CAASPP training on Smarter

Balanced components, and sometimes these trainings are made available to teachers. However, many schools have not had the time to pass along information to all their staff. Some teachers had not tried logging on to the IA Reporting System since the many enhancements to it were launched in 2018–19. Most teachers in the small study sample had not explored the DL, often because they found logging on confusing or because they felt they had sufficient resources already. Teachers in the study who tried the DL noted frequently that navigation was difficult and time consuming, though some of these teachers may have been referring to earlier versions of the system before it was enhanced.

Recommendation 5: The CDE should seek ways to streamline or provide additional guidance on rostering within the IA Reporting System, including recommendations regarding what access LEAs should be providing to their teachers. Some CAASPP coordinators found the CAASPP rostering process to be cumbersome, and for one LEA there was confusion in 2017–18 that resulted in teachers not having student-level results. In addition, some teachers would like more access than they are currently provided by their school or LEA. Accessibility of IA report features at the educator level is dependent upon the creation of rosters by the coordinator. Teachers may benefit if their CAASPP coordinators are given more direction regarding what level of access they should provide their teachers.

California Science Test Alignment Study

Overview

HumRRO's alignment studies for the CAASPP evaluation are designed to gather evidence to help demonstrate the validity of intended interpretations and uses of the assessment scores. The CAST alignment study will evaluate how well the 2019 test items fully sample the construct represented by the associated content standards, the California Next Generation Science Standards (CA NGSS). That is, the alignment study will indicate whether the CAST effectively measures what it is intended to measure.

The CAST is a computer-based, fixed-form (non-adaptive) assessment administered to students in grades five, eight, and one time in high school (i.e., grade ten, eleven, or twelve). The CAST was administered operationally for the first time in January–July 2019. The assessment included three segments: Segment A, consisting of discrete items (e.g., selected-response, short constructed-response, technology enhanced items); Segment B, consisting of two performance tasks; and Segment C, consisting of field test items (discrete) or a performance task. The CAST alignment study is based on Segments A and B only because only Segments A and B were administered operationally in 2019.

HumRRO modified traditional alignment methods to account for CAST's structure and design, a process in keeping with best practices in test validation that facilitates using alignment study results in an overall validity argument. This modified process also supports federal peer review goals. The CAST was developed to measure student achievement in the CA NGSS performance expectations (PEs), which are assessable

statements of what students should know and be able to do. The three major components of the CA NGSS—Disciplinary Core Ideas (DCIs), Crosscutting Concepts (CCCs), and Science and Engineering Practices (SEPs)—are the dimensions that operationalize the PEs. Developing tests and test items that adequately sample such complex and integrated content as the CA NGSS is especially challenging. When an item measures a single standard or concept, the alignment process is relatively straightforward. However, test development and alignment become more complex when standards are designed as interactions among statements about the three dimensions of standards. HumRRO responded to these challenges with our study design, described in depth in the 2018–20 CAASPP Evaluation Plan.

To allow the CAST to address the full breadth of the CA NGSS, it was designed to rotate content across a three-year span, such that different content from the CA NGSS is sampled each year. Because HumRRO is conducting this alignment study after the first operational year of testing, it will not be possible to evaluate how well CAST addresses the full breadth of the content standards over three years. However, HumRRO will use the initial year's data to estimate whether one administration can address roughly one-third of the intended PEs. If so, the three-year rotation is feasible as a sampling plan for addressing the full breadth of the CA NGSS.

The next sections describe activities conducted to date. Chapter 3 presents the study in greater depth, including the research questions and methods and activities conducted during 2018–19. Upon conclusion of the study in 2020, HumRRO will provide responses to the research questions in an alignment study report, which should guide future item development and provide validity evidence for the CAST suitable for submission for federal peer review under ESSA.

Progress Made to Date

Evaluation of CAST Contractor Documentation

HumRRO conducted an initial review of contractor documentation to evaluate how alignment issues were considered during test development. This review was guided by the *Standards for Educational and Psychological Testing* (APA, AERA, NCME, 2014, hereafter referred to as the *Test Standards*). The *Test Standards* describe requirements for developing, reviewing, and piloting test items as well as requirements for documenting the processes used. In a draft report submitted to the CDE on January 24, 2019, HumRRO summarized preliminary findings based on the initial review of CAST documentation provided by the CDE and the testing contractor, ETS. Following that preliminary report, HumRRO requested additional documentation, which was provided by the testing contractor. HumRRO will review and evaluate the additional documentation and include these findings in the 2020 CAST alignment study report.

CAST Alignment Criteria

HumRRO developed alignment criteria for the CAST study based on the Webb alignment method (1997, 1999, 2005), which includes the following four indicators:

categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance-of-knowledge representation. While it was not appropriate to implement Webb's alignment method for the CAST study, mainly because of the multidimensional nature of the content standards and the way the content is sampled across years, HumRRO did use Webb's criteria to modify the method and develop criteria to judge alignment of the CAST.

HumRRO's draft criteria were reviewed by staff from the National Center for Improvement in Educational Assessment (Center for Assessment) in January 2019. Reviewers offered several comments to clarify how the criteria would be communicated and operationalized for the study. The criteria were presented to the CDE and its CAASPP Technical Advisory Group (TAG) in February 2019 and finalized prior to the CAST alignment workshop.

Panel Evaluation of CAST Item Content

HumRRO worked collaboratively with the CDE to recruit and select a group of 18 educators to serve on three CAST alignment review panels (one elementary, one middle, and one high school panel). Panelists were very familiar with the CA NGSS and were required to have at least three years of experience as California educators.

HumRRO conducted a two-day CAST Alignment Study Workshop in the Sacramento area on February 28–March 1, 2019. During the workshop, panels of educators evaluated how well each CAST item assessed the CA NGSS. The panelists made ratings regarding what content standards the items assessed, accounting for the three-dimensional nature of the CA NGSS. Panelists also rated each item according to its cognitive complexity requirements. Panels discussed discrepant ratings and reached clear consensus or near-consensus when they disagreed about ratings. HumRRO is analyzing the CAST alignment workshop data and will include outcomes in the 2020 CAST alignment study report.

California Alternate Assessment for Science Alignment Study

Overview

The CAA for Science alignment study aims to provide validity evidence as a measure of science achievement for the population of students for which the assessment was designed. The CAA for Science is a fixed-form (non-adaptive) assessment administered to students with the most significant cognitive disabilities in accordance with each student's active individualized education program (IEP). This group makes up approximately 1% of the total population of students in California. The CAA for Science is given in grades five, eight, and high school as three separate operational test sessions. Each session consists of one fixed-form embedded performance task (i.e., one for life sciences, one for physical sciences, and one for earth and space sciences). Each performance task includes 10 items and is intended to function similarly to an "end-of-instruction" test rather than an "end-of-year" summative assessment. The test is

structured such that the first five items measure one standard and the last five items measure a second standard. The CAA will be administered operationally for the first time September 2019–July 2020.

The alignment study research questions and method were designed specifically to address the structure and design of the CAA for Science and the results to be reported. This study will focus on links between the Science Core Content Connectors (alternate achievement standards) and the test forms and items developed to assess them. The Science Core Content Connectors (hereafter referred to as "Science Connectors") are based on the performance expectations of the CA NGSS, which also define the science construct(s) to be measured.

The next sections describe activities conducted to date. Chapter 4 presents the study in greater depth, including the research questions and methods and activities conducted during 2018–19. Upon conclusion of the study in 2020, HumRRO will provide responses to the research questions in an alignment study report, which should guide future item development and provide validity evidence for the CAA suitable for submission for federal peer review under ESSA.

Progress Made to Date

Coordination with CAA for Science Test Contractor and the CDE

HumRRO's project manager and the CAA for Science Alignment Study director met with staff from the testing contractor (ETS) for CAA for Science and CDE staff to coordinate study activities. Meeting participants discussed (a) HumRRO's plans for data collection, (b) CAA for Science assessment materials (e.g., online test content, Directions for Administration, planning guides), (c) documentation needed from ETS and CDE, (d) estimates of dates when files would be available to HumRRO from ETS, and (e) panelist recruitment. Based on when ETS could provide all materials and process support for the alignment workshop, HumRRO scheduled the alignment workshop for November 2019.

Evaluation of CAA for Science Contractor Documentation

Similar to the initial steps for the CAST Alignment Study, HumRRO requested contractor documentation for the CAA for Science and began reviewing the first submittals to evaluate how alignment issues were considered during test development. As for the CAST Alignment Study, the review of CAA for Science documentation is guided by the *Test Standards*. After initial review, HumRRO will produce a preliminary report of findings and identify any gaps in the documentation. HumRRO will follow up with the testing contractor to ask questions and request additional documentation until all *Test Standards* are independently rated. HumRRO will include findings from the CAA for Science alignment workshop in the 2020 CAA for Science alignment study report.

Preparing for the CAA for Science Alignment Workshop

HumRRO is working collaboratively with the CDE contract monitor to recruit 18 educators to serve on three CAA for Science alignment review panels (one grade five, one grade eight, and one high school panel). Panelists are required to have a bachelor's degree and experience as a California teacher, to include experience working with severely cognitively disabled students or students with mild to moderate disabilities. Ideally, most panelists will also have familiarity with the CA NGSS and the Science Connectors.

HumRRO secured a venue in the Sacramento area for the two-day workshop and began arranging lodging and travel for confirmed panelists and planning for all necessary materials, processes, and equipment.

Summary and Next Steps

The first year of HumRRO's Impact Case Study provided an in-depth look at how a modest number of diverse LEAs and schools are implementing Smarter Balanced components, especially the interim assessments. Overall findings indicate the IABs, which are high quality, CCSS-aligned online assessments, are still mainly used to prepare students for the rigor and format of the summative assessments. However, for the general education population of students, teachers are increasingly using IABs, along with other measures of student progress, in creative and effective ways to assist with instructional decisions, plans, and goals. The CDE and its vendors continue to make substantive improvements to the various components supporting the Smarter Balanced assessments; however, not all LEAs and schools are keeping current on the training and resources available to understand and use the enhanced features. We fully support the CDE's continued efforts to implement solutions to areas identified for improvement, internally and by our independent evaluation, as the CAASPP System matures.

For the second year of the Impact Case Study, HumRRO will continue to focus on the Smarter Balanced components of the CAASPP System. For the most part, the same data collection activities will be conducted, although with a different group of LEAs and schools. HumRRO's research will aim to learn how best to support teachers' awareness of the full range of CAASPP components and what kinds of experiences teachers need to be able to implement the interim assessments and DL for instructional purposes. HumRRO's progress on the CAST and CAA for Science alignment studies is on track for concluding the studies and producing their respective technical reports in 2020.

This page is intentionally blank.

Chapter 1: Introduction

Background

The California Assessment of Student Performance and Progress (CAASPP) System, launched in 2014, was intended to assist teachers, administrators, students, and parents by promoting high-quality teaching and learning using a variety of assessment approaches and item types. The statewide student assessments monitor progress in implementing effective instruction aligned with the Common Core State Standards (CCSS) for English language arts/literacy (ELA) and mathematics and the California Next Generation Science Standards (CA NGSS). The Smarter Balanced ELA and mathematics tests have been operational since 2014, and the California Alternate Assessments in ELA and mathematics have been operational since 2016. The California Science Test (CAST) became operational in spring 2019, and the California Alternate Assessment in Science (CAA Science) will become operational during the 2019–20 school year. The CAASPP System also includes an optional Spanish reading language arts test, the California Spanish Assessment (CSA), which became operational in 2019. These assessments aim to shift the focus away from accountability toward a comprehensive plan for promoting high-quality teaching and learning for all students, including students with disabilities (SWDs) and English learners (ELs). The CAASPP System represents a substantial financial investment by the state as well as a significant investment of educator and student time.

California *Education Code* (*EC*) Section 60649 requires the independent evaluation of the CAASPP System, stating that "evaluation activities may include a variety of internal and external studies such as validity studies, alignment studies, and studies evaluating test fairness, testing accommodations, testing policies, and reporting procedures, and consequential validity studies specific to pupil populations such as English learners (ELs) and pupils with disabilities." The law requires development of a plan to assess independent evaluation activities, and it prohibits duplication of studies conducted as part of a federal peer review process or by California Department of Education (CDE) assessment contractors.

The Human Resources Research Organization (HumRRO) served as the first CAASPP System evaluator from 2015–18. Copies of our annual and comprehensive final reports are available on the California Department of Education (CDE) Web page (https://www.cde.ca.gov/ta/tg/ca/caaspprptstudies.asp).

The CDE awarded the contract for the 2018–20 independent evaluation of the CAASPP System to HumRRO in July 2018. The current contract calls for annual evaluation reports that summarize all work completed during the previous year, stand-alone reports for individual research studies, and a comprehensive final report. Within a few months of the award, HumRRO submitted to the CDE the first required annual evaluation report (Hardoin, Thacker, Dvorak, Becker, 2018). That report's core contents included the 2018–20 Evaluation Plan, which described the design of three research studies approved by the CDE and scheduled within the contract period. The present report is

the second annual report, and it describes activities conducted and results obtained to date from the 2018–19 studies. The third annual report will describe results from the 2019–20 studies. A Comprehensive Final Evaluation Report will be delivered in 2020 and will include evaluation findings from each of the three annual reports (2018, 2019, and 2020).

An ongoing evaluation is important to support the goal of continuous improvement to help California achieve the intended return on its investment in the CAASPP System. The evaluation can provide evidence to demonstrate the validity of intended interpretations of test scores used as measures of student learning relative to targeted content standards, and it can offer recommendations for potentially improving alignment between what an assessment is intended to measure and what it actually measures. The evaluation can also provide insight into how CAASPP results are used to improve instruction at the student, classroom, school, local educational agency (LEA), and statewide levels.

2018–20 Evaluation Plan Goals and Timeline

As context for this year's report on evaluation activities, table 1.1 gives an overview of the goals and schedule for each research study included in the 2018–20 Evaluation Plan. HumRRO developed the plan with guidance from the CDE and input from the CAASPP Technical Advisory Group (TAG). Each research study was designed to provide information about how well specific parts of the CAASPP System as delivered meet the intended goals of the program expressed in the CAASPP System theory of action. The plan in its entirety is available in the 2018 Independent Evaluation Report (https://www.cde.ca.gov/ta/tg/ca/documents/caaspp18evalrpt.pdf).

Table 1.1 Overall Goals and Schedule for Each 2018–20 Evaluation Study

Study Title	Goals and Schedule
Impact Case Study	• Collaborate with and gather extensive qualitative data (case studies) from a small sample of schools and LEAs, purposefully selected based on their use of CAASPP components and resources. The small sample will aim to broadly represent the diversity of the state with respect to geographic location, academic achievement, and size (student enrollment), as well as student population characteristics (i.e., socioeconomic disadvantage and EL status). • Investigate the context and various approaches used by the small sample of schools and LEAs to implement and integrate the CAASPP System components to inform instruction and improve student learning.

Table 1.1 (cont.)

Study Title	Goals and Schedule
Impact Case Study *(cont.)*	• Impact Case Study reports will each describe in detail one school year's findings of the studied LEAs' and schools' use of CAASPP components and their impacts on instruction and student learning. The report will document in detail the local context for each case study. • Conduct year one data collection activities with initial set of LEAs and schools in 2018–19. • Complete year one data analysis and develop stand-alone year one report in 2019. • Conduct year two data collection activities with second set of LEAs and schools in 2019–20. • Complete year two data analysis and develop stand-alone year two report in 2020.
CAST Alignment Study	• Evaluate the degree of alignment between the CAST test items and test forms with the CA NGSS. • CAST Alignment Study Report should guide future item development and provide validity evidence suitable for submission for federal peer review under the Every Student Succeeds Act (ESSA). • Conduct data collection activities in 2018–19. • Complete data analysis and develop stand-alone report in 2019–20.
CAA for Science Alignment Study	• Evaluate the degree of alignment between the CAA for Science test items and test forms with the Science Connectors and Focal Knowledge, Skills, and Abilities (FKSAs) derived from the CA NGSS. • CAA for Science Alignment Study Report should guide future item development and provide validity evidence suitable for submission for federal peer review under ESSA. • Conduct data collection activities in 2019–20. • Complete data analysis and develop stand-alone report in 2019–20.

Implementing the 2018–20 Evaluation Plan

A summary list of key activities and time frames for implementing the 2018–20 Evaluation Plan is presented in table 1.2, along with a status of the work as of June 30, 2019.

Table 1.2 Schedule and Status of Evaluation Activities for 2018–20

Activity	Time Frame	Status
Orientation Meeting with CDE staff: In-person meeting to review all tasks and project timeline and to address questions and concerns.	July 2018	Completed
Management Meetings with CDE staff: Biweekly calls to discuss progress, plans, and issues.	July 2018–December 2020	In Progress
State Board of Education (SBE) Meetings: Meet with SBE staff and provide presentations at Board meetings.	As requested, up to two times annually, July 2018–December 2020	To Be Scheduled
TAG Meetings: Meet with and provide presentations, including detailed designs, review of progress on studies, preliminary findings from studies, and Evaluation Plan updates.	Three times annually, July 2018–December 2020	In Progress
CAASPP Contractor Annual Planning Meeting: Attend meeting to learn of planned updates to the system, concerns, processes, scope, and schedule.	Annually, July 2018–June 2020	In Progress
Conduct the CAST Alignment Study and deliver a stand-alone study report.	July 2018–June 2020	In Progress
Conduct the Impact Case Study and deliver two stand-alone study reports.	Annually, July 2019–December 2020	In Progress
Conduct the CAA for Science Alignment Study and deliver a stand-alone study report.	July 2019–June 2020	In Progress
Develop and deliver annual report.	Annually, July 2018–December 2020	In Progress
Develop and deliver final comprehensive report.	July–December 2020	Scheduled
Maintain comprehensive plan and schedule for project activities and deliverables.	July 2018–December 2020	In Progress
Submit monthly written progress reports to describe evaluation progress, plans, and issues.	July 2018–December 2020	In Progress

During 2018–19, HumRRO conducted two critical activities prior to or in concurrence with the two specific research studies: (a) observation of CAASPP Smarter Balanced educator training sessions, and (b) internal training of HumRRO's CAASPP evaluation project team on the security and confidentiality procedures for handling evaluation data.

Background Research on Updated CAASPP System

During the first CAASPP evaluation contract, HumRRO researchers reviewed resources to build knowledge of the Smarter Balanced components and CAAs for ELA and mathematics. We reviewed publicly available online information, attended educator training sessions, and obtained access to other resources (e.g., the Smarter Balanced Digital Library, weekly *CAASPP Update* emails) to understand how the components were presented to California teachers, administrators, and district staff. HumRRO's data collection activities for the first evaluation contract ended in June 2017. When the second contract began, HumRRO researchers supplemented their foundational knowledge with updated information about the CAASPP System. This included reviewing SBE meeting minutes and subscribing to *Assessment Spotlight,* CDE's renamed weekly email to educators from kindergarten to grade twelve. Launched on July 5, 2018, this publication includes information about CAASPP as well as the *English Language Proficiency Assessments for California (ELPAC).*

For the Impact Case Study, HumRRO developed data collection instruments and an analysis plan. Both these processes required current knowledge of CDE-hosted LEA and school staff training on the Smarter Balanced summative and interim assessments and reporting systems, as well as the Digital Library. HumRRO found information about these sessions was readily available in the "Training" tab of the CAASPP portal, which describes goals of in-person professional development opportunities and provides links to videos and archived webcasts of sessions and materials. Some sessions are offered multiple times and conducted at different locations in the state, and the Training tab organizes the opportunities by month and target audience (e.g., classroom teacher, CAASPP coordinator) to give an at-a-glance view of summer and upcoming school year offerings.

HumRRO's project manager observed two CAASPP training sessions, actively attended to the content of presentations, studied the materials provided, and engaged with other participants during small group activities. Below are brief overviews of the observed sessions:

- *2018–19 CAASPP Institute – Implementing the Smarter Balanced System of Assessments* (October 22–23, 2019; San Diego, California). This two-day workshop was conducted by the Sacramento County Office of Education (SCOE) in partnership with the CDE. Participants were directed to 13 folders of materials accessible online prior to the workshop. Some folders included handouts used during the workshop and others included resources for future reference. In addition to large group sessions covering the basics of each Smarter Balanced component, break-out sessions involved accessing online interim assessment (IA) results in the IA Reporting System and resources in the Digital Library. The break-out sessions were particularly engaging and eye opening to those who had not ever accessed these resources before or had not accessed them since significant enhancements were made to the features, functions, and filters of the online resources.

- *The Results Are In…Now What? Analyzing Assessment Results to Inform Teaching and Learning Workshop* (May 31, 2019, Downey, California). This one-day workshop was conducted by WestEd, and LEAs were encouraged to attend as a team (e.g., LEA CAASPP coordinators, professional development staff, and curriculum specialists). Links to handouts and resources were provided two weeks before the workshop, and participants were told to bring a laptop and their Online Reporting System (ORS) log-on information to maximize the day's learning opportunities. Presenters focused on building assessment literacy, what to do with results of different assessments (formative, interim, summative) and the different types of data they produce, and how to avoid errors in using results (e.g., using them for purposes for which they aren't intended). The presenter reviewed a data analysis protocol but emphasized there are many approaches to working with data to improve student learning. HumRRO observed table teams actively using ORS to review preliminary 2019 Smarter Balanced results, comparing them to the prior year's, and discussing past and current programs, policies, and practices to help them understand their students' scores.

CDE's training materials for the Smarter Balanced components emphasize the potential to impact teaching and learning when the CAASPP System tools are used in conjunction with each other. Additionally, CDE training materials highlight the critical purpose of student assessment: to gather evidence to make informed and appropriate instructional, policy, and programmatic decisions based on data. Figure 1.1 includes two diagrams from a 2018–19 CAASPP Institute presentation, Integrating the CAASPP Tools to Create a Process of Improvement (CDE, 2018b). The left diagram on the left is a framework for thinking about how curriculum, instruction, and assessment fit into the cyclical process of teaching and learning. The diagram on the right illustrates the interconnectedness of the Smarter Balanced components, which align with the framework to produce continuous improvement in student learning. While encouraging educators to use all the free CAASPP components, the training states there is no single best way to maximize the information provided by the CAASPP components. The training also recognizes CAASPP components must be applied to suit the context of a classroom, school, or district, along with other formative processes vital to the teaching and learning cycle. See Appendix F for a detailed description of the figure.

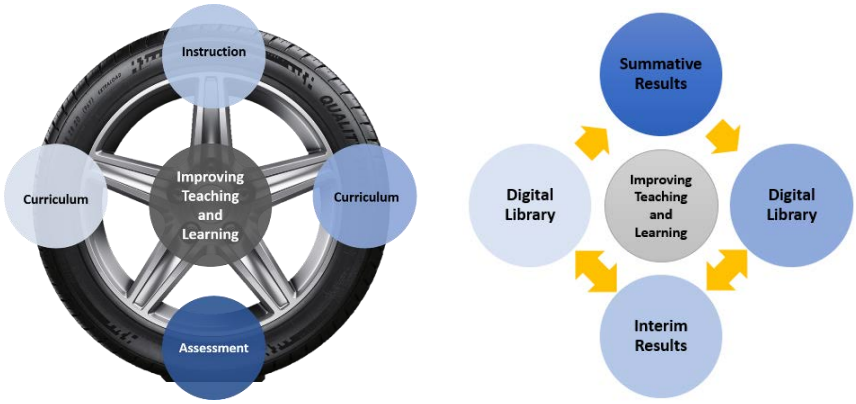


Figure 1.1. Framework for an integrated set of assessments, from 2018–19 CAASPP Institute materials.

Safeguarding Confidential Data

HumRRO fully understands the importance of adhering to policies that protect and monitor access to sensitive information, such as confidential test materials and focus group and interview data, while carrying out the independent evaluation activities. HumRRO researchers are cognizant of federal policies such as the Federal Educational Rights and Privacy Act (FERPA) as well as policies pertaining to governmental agencies in California and those specific to the CDE.

For the CAASPP evaluation, staff security program training focused on three key areas: (a) proper administration of non-disclosure agreements and implementation of the "need to know" principle for all personnel working on the contract; (b) comprehensive training on specific security requirements related to HumRRO's CAASPP work, including but not limited to, specific data security and incident report procedures; and (c) clear explanation of pertinent laws and regulations governing—and the procedures related to protecting—the safeguarding of certain types of information relevant to the contract. Taken together, these areas of our security program ensure all procedures are administered in an efficient and effective manner.

Organization and Contents of the 2019 Evaluation Report

The remaining chapters of this report describe progress made during 2018–19 in each research study listed in table 1.1.

- Chapter 2, *Impact on Instruction and Student Learning Case Study*, presents HumRRO's methods and data collection activities conducted to study the use of CAASPP Smarter Balanced components (i.e., summative and interim assessments and the Digital Library) by a small number of LEAs and a small subset of each LEA's schools. The goals of the study were to learn how the CAASPP System impacts ELA and mathematics instruction and student learning by collecting and analyzing extensive qualitative data about the use of the components in the specific context of the LEAs and schools. HumRRO collected data via in-person focus groups and interviews, monthly email polling, end-of-year web-based focus groups, and school-led student focus groups. The chapter provides, for each research question, the overarching themes and unique aspects discovered in the use of Smarter Balanced components across the LEAs. The chapter concludes with best practices and recommendations for effective use of the Smarter Balanced components.

- Chapter 3, *California Science Test (CAST) Alignment Study*, presents the research questions and a summary of the methods and data collection activities completed to date to investigate the alignment of CAST to the CA NGSS. The chapter describes HumRRO's implementation of the alignment workshop plan in coordination with support and documentation received from CDE's CAST test contractor. The chapter also presents the alignment acceptability criteria HumRRO developed for this study. HumRRO will present outcomes of analysis of the alignment data and evaluation of CAST contractor documentation in a stand-alone report in 2020.

- Chapter 4, *California Alternate Assessment (CAA) for Science Alignment Study*, presents progress HumRRO has made to date to begin investigating the alignment of CAA for Science to the Science Connectors and FKSAs derived from the CA NGSS. HumRRO will conduct the study during 2019–20 and present outcomes in a stand-alone report in 2020.

Chapter 2: Impact on Instruction and Student Learning Case Study

The two-year Impact on Instruction and Student Learning Case Study (hereafter, Impact Case Study) uses a case study approach to deeply investigate and produce a richly detailed summary of the CAASPP System's impact in a modest number of local educational agencies (LEAs) and schools. During the 2018–19 school year, the first year of the study, HumRRO collaborated with seven LEAs, including one direct-funded charter school. The primary goal of the study is to elicit concrete examples of how and why specific CAASPP components (i.e., Smarter Balanced components for English language arts/literacy [ELA] and mathematics) were used, their impact on instruction and student learning, as well as the perceived benefits, strengths, and challenges of using the components.

Creswell (1998) described a case study as an appropriate research approach when one is interested in the in-depth study of a "case" bounded in time or place. Patton (2015) noted that a "case" can be many different things, depending on the focus and field of study. Moss and Haertel (2016) use the label "Small N or Comparative Case Studies" (CCS) for studies with "more than one case, but typically fewer than fifty, purposively chosen to illuminate the question or phenomenon of interest. Typically, cases are chosen so as to contrast with respect to some set of key features. In CCS, within-case analyses are supplemented by cross-case comparisons, which help to support generalization."

For this study, a case was defined as an LEA that had fully implemented the CAASPP System in 2017–18 and planned to continue implementation during the study year, 2018–19 (see description in Selection of LEA Cases). To conduct a case study, one should gather a large amount of data to provide an in-depth picture of the "case" (Creswell, 1998). Like other forms of qualitative research, case studies tend to rely on use of inductive reasoning, rather than beginning with specific hypotheses (Creswell & Plano-Clark 2007). Consistent with these approaches, HumRRO's study methods relied on inductive reasoning guided by a set of research questions. HumRRO incorporated multiple types of data collection, as described further in this chapter, to provide an in-depth look at the implementation of CAASPP for a selection of LEAs and a sample of their schools.

The candor and thoughtfulness of study participants' responses to questions during all phases of data collection were the foundation of this study. HumRRO researchers express our deep gratitude for the time, collaboration, and contributions made by LEA and school staff to this important work.

This first section of this chapter describes the CAASPP components studied. The second section presents an abbreviated version of the study design and describes the recruitment and selection of LEAs and their associated schools. The detailed design of the Impact Case Study is included in the 2018–20 CAASPP Evaluation Plan, which is presented in the publicly available *2018 CAASPP Evaluation Report*. The current report provides briefer descriptions of each aspect of the study design, including modifications made during implementation of the study, to give context for the reporting of findings.

The next section of this chapter presents a summary of findings from each of the LEAs and its sample of schools. Each summary includes (a) an overview of the context of the LEA and its schools, (b) a summary of findings about usage of each CAASPP component, (c) identification of best practices in the use of CAASPP components, and (d) several recommendations for improvement to achieve more effective use of the components to impact instruction and student learning.

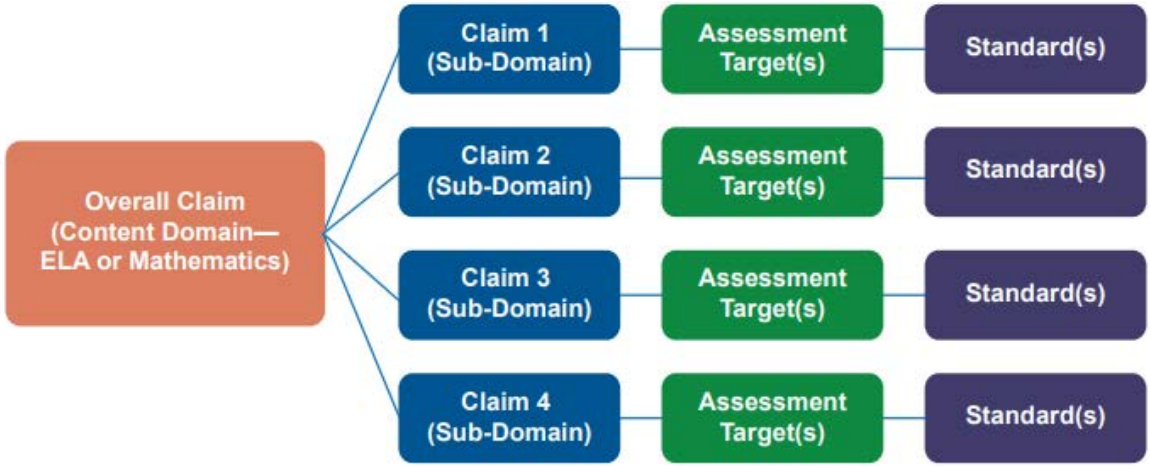
The LEA-specific section of this chapter is followed by general findings regarding CAASPP component use across all the LEAs studied this first year, organized by the research questions of the study. This section includes HumRRO's evaluation of contextual implications, common experiences, best practices, and challenges. The outcomes of year one of the Impact Case Study will inform CDE about successes as well as obstacles and suggest where potential future improvements can be made to increase the CAASPP System's intended utility to positively impact classroom instruction and student learning. The chapter concludes with a summary list of best practices and recommendations for further improvements, and a brief overview of next steps for year two of the study.

CAASPP Smarter Balanced Components

The CAASPP System comprises multiple components intended to measure student performance and progress and serve as tools for increasing student learning in the classroom. This Impact Case Study focused only on the CAASPP Smarter Balanced components for ELA and mathematics, all of which were intentionally designed to align to the content and rigor of the Common Core State Standards. Figure 2.1, from the CDE's Teacher Guide to the Smarter Balanced Summative Assessments (2016), illustrates how overall domain claims (most general level), sub-domain claims, assessment targets, and standards (most specific level) form a hierarchy whose different levels guide test development and contribute to analysis and understanding of different types of Smarter Balanced scores. There are four sub-domain claims for ELA (reading, listening, writing, and research/inquiry) and four sub-domain claims for mathematics (concepts and procedures, problem solving, modeling and data analysis, and communicating and reasoning). Test results for mathematics collapse two of the mathematics claims (problem solving and modeling and data analysis) into one score reporting category.

The Smarter Balanced Summative Assessments for grades three through eight and eleven are the only Smarter Balanced component required for use in a standardized manner by all California public schools, including charter schools. The end-of-year summative assessments are used as the state accountability tests for ELA and mathematics and are delivered by computer. Each assessment includes a computer adaptive test and a performance task (PT). The computer adaptive test includes a variety of item types such as selected response, constructed response, table, fill-in, and graphing, while the PTs are extended activities that measure integration of knowledge and skills across multiple standards. The CDE provides access to aggregate results from the Smarter Balanced Summative Assessments on its public website (e.g., for students, parents, educators, researchers). Individual student reports are available only to LEA CAASPP coordinators and school test site coordinators and to parents or guardians and

may be obtained only from the schools and districts where students were tested. LEAs and schools have access to a variety of score reports for their students in the Online Reporting System (ORS), and they may also download data from that system.



See Appendix F for a detailed description of the figure.

Figure 2.1. Hierarchy of Smarter Balanced item development

The CAASPP System includes Smarter Balanced tools—Interim Assessments and the Digital Library—that are not required but are available to California schools throughout the school year. The CAASPP System included two main types of Smarter Balanced Interim Assessments in the 2018–19 school year, Interim Assessment Blocks (IABs) and Interim Comprehensive Assessments (ICAs). IABs are brief assessments focused on small sets of assessment targets and provide detailed results for instructional purposes. ICAs are built using the same blueprints as the summative assessments and provide results on the same scale. All ICAs and some IABs include constructed response items that require local hand scoring. The Interim Assessment Reporting System (IA Reporting System) provides results for IABs and ICAs. IA results include group-level analysis (average scale score and distribution of scores across performance levels), group item-level analysis (proportion of students at each score point and item information, including item difficulty and the claim, target, and standard assessed), student-level analysis (item information, including depth of knowledge, and student responses), key and distractor analysis, and writing trait scores. Depending on how the IA was administered, results can be used by teachers "to identify students who have a strong grasp of the material and need enrichment activities to support expansion of their skills; group students by knowledge/skill level for differentiated instruction; and pinpoint areas to emphasize during classroom instruction" (Smarter Balanced Assessment Consortium, 2019)

The Smarter Balanced Digital Library (DL) provides instructional resources for educators to use during daily instruction in support of the formative assessment process. Individual resources can be accessed through a search by subject, grade level, specific CCSS or target, intended student population (e.g., English learners, students with disabilities), and other characteristics. Alternatively, educators can access playlists, which are collections of DL resources focusing on similar content and

organized by progressions of skills or understandings. Playlists and individual resources are also accessible through links in the IA Reporting System. This functionality allows teachers to be connected directly to resources in the DL that target their students' needs. The DL also provides professional learning resources with teaching strategies.

The intended purposes of the Smarter Balanced summative assessments, interim assessments, and formative assessment resources are specifically described in the theory of action for the CAASPP program (see the *CAASPP 2018 Independent Evaluation Report,* https://www.cde.ca.gov/ta/tg/ca/documents/caaspp18evalrpt.pdf, Appendix A). There are supplemental CAASPP Smarter Balanced resources available to California educators to assist with using these CAASPP components, such as in-person training workshops and archived workshop presentations, webcasts, online manuals, and videos. Additionally, Smarter Balanced ELA and mathematics practice tests are available online to prepare students and teachers for the summative assessments.

Study Design and Selection of LEA Cases

Research Questions

The Impact Case Study addresses 13 key research questions pertaining to the CAASPP components of interest. Questions are organized into three general areas: (a) contextual questions and those pertaining to the full suite of Smarter Balanced components in the CAASPP System, the Summative Assessments, Interim Assessments, and Digital Library of formative tools; (b) questions related only to the Smarter Balanced Summative Assessments; and (c) questions related to the Smarter Balanced Interim Assessments and Digital Library resources. Table 2.1 presents the research questions and the components they address. These questions serve as the organizing structure for the presentation of the findings. HumRRO's investigation of answers to the research questions was limited to data collection from participating staff from the small sample of selected LEAs and their few selected schools.

Contextual conditions are important influencers of the implementation of policies and practices, as noted in a recent literature review of interventions to support educators' use of data to guide decision making and practices (Marsh, 2012). Contextual conditions can be tied directly to use of data, such as the "capacity of the intervener" (e.g., guide or deliverer of training for data interpretation) and data properties (e.g., ease of interpreting outcomes of multiple measures). Broader contextual conditions include "leadership, organizational structure, time, [and] policy," as well as "interpersonal relationships and belief and knowledge."

HumRRO explored LEA and school context in terms of many factors—student demographic characteristics; academic achievement in ELA and mathematics; teacher turnover; class scheduling considerations; available curricular, technological, and other resources; professional development opportunities; and the role of professional learning communities of all types. For the purpose of this evaluation, the acronym PLC is used as an umbrella term for organized small groups of teachers who meet regularly to collaboratively develop practice-based professional learning.

Table 2.1 CAASPP Components and Impact Case Study Research Questions

CAASPP Components Addressed	Research Questions for Sampled LEAs and Schools
Summative, IABs and ICAs, DL	1. What are the characteristics and contexts of sampled schools/LEAs that have implemented the full suite of Smarter Balanced components?
Summative, IABs and ICAs, DL	2. How does implementation of Smarter Balanced components vary across schools/LEAs? What instructions and supports are provided to educators for implementing the components?
Summative, IABs and ICAs, DL	3. What aspects of Smarter Balanced components are perceived as most beneficial for improving classroom instruction and student learning across schools/LEAs?
Summative, IABs and ICAs, DL	4. What changes to the components and supporting resources do LEA and school staff believe would improve support for their use to improve classroom instruction and student learning?
Summative, IABs and ICAs, DL	5. How do educators/schools/LEAs use and integrate results from the summative, interim, and formative assessment resources for each content domain with each other and with other measures to enhance classroom instruction and student learning? What challenges are faced and how are they overcome?
Summative, IABs and ICAs, DL	6. How do students from schools that use the full suite of components perceive classroom opportunities to learn about summative assessment item types and topics?
Only Summative assessments	7. How do educators/schools/LEAs use summative assessment data to inform classroom instruction and make decisions?
Only IABs, ICAs, and DL	8. What interim assessments are used for ELA/literacy and mathematics for schools/LEAs that have implemented the full CAASPP System, and at what grade levels and frequency?
Only IABs, ICAs, and DL	9. What decision-making processes are used by educators/schools/ LEAs to determine what interim assessments to use, who should administer them, and how frequently?
Only IABs, ICAs, and DL	10. To what extent have educators/schools/LEAs incorporated IABs into their classes? What, if any, classroom assessments have been replaced in the process? Why, and what are the implications?
Only IABs, ICAs, and DL	11. How do educators/schools/LEAs use information from ELA/literacy and mathematics interim assessments to track individual student progress and/or inform classroom instruction?
Only IABs, ICAs, and DL	12. How is information on student/school/LEA performance on interim assessments used at the school/LEA level to determine the effectiveness of practices and curricular materials for teaching the targeted standards?
DL	13. How is the DL used to improve classroom instruction?

LEA Sample

For the Impact Case Study, HumRRO's goal was to identify and recruit six LEAs (including one charter school) that used all three CAASPP Smarter Balanced components (summative assessments, IAs, and DL) according to criteria developed jointly between HumRRO and CDE (Hardoin, Thacker, Dvorak, Becker, 2018):

> *These LEAs should have demonstrated during the 2017–18 school year at least a "modest threshold" of use of both of the optional Smarter Balanced CAASPP components (a) IAs, with or without ICAs and hand scoring, and (b) the Instructional Resources of the Digital Library, with or without use of Professional Learning resources and Playlist resources. "Modest threshold" means a sufficient amount of use beyond simply investigating system features and will be defined based on Digital Library log-on data and interim assessment data provided to HumRRO. Eligible LEAs need not be the heaviest users in the state.*

After a review of 2017–18 school year IA usage data and DL log-on data, HumRRO identified the thresholds for LEA participation in the first year of the study and received CDE's approval for these eligibility criteria. HumRRO's cut point for IA usage required LEAs to have administered at least 300 IABs in ELA, mathematics, or both domains, during 2017–18. No requirement was established for ICA administration, as ICA usage was much less extensive than IAB usage. Due to the lack of extensive use of the DL in any LEA, HumRRO established a very modest threshold of at least 10 users in the LEA who logged on to the DL during 2017–18. After using these criteria to prescreen potential LEAs, HumRRO administered the 2018 Eligibility Survey (see Appendix D in the *2018 Independent Evaluation Report*) to all LEAs that met the minimum requirements. HumRRO administered the brief survey to further refine the set of eligible LEAs by collecting additional information about their CAASPP involvement, school characteristics, and willingness to participate in the Impact Case Study. HumRRO sent an invitation to complete the online survey to LEA CAASPP Coordinators. Table 2.2 summarizes survey respondents by LEA type (overall 25% response rate) and interest in participating in the study.

To choose cases from the eligible LEAs, HumRRO implemented the sampling plan outlined in the *2018 CAASPP Evaluation Report*. The goal was to identify LEAs that would very broadly represent the diversity of the state in terms of geographic region, student enrollment and demographics, and academic achievement. For each LEA, HumRRO sought to include one elementary school, one middle school, and one high school. HumRRO did not seek a representative sample of schools from each LEA, but rather identified a sample of schools that were strong implementers of the Smarter Balanced components. HumRRO communicated with 21 LEAs to reach the target number of cases. By December 2019, HumRRO had commitment from six LEAs. In January 2019 however, one of the LEAs declined to participate further because of the time required for school staff to contribute to ongoing data collection. HumRRO recruited a seventh replacement LEA. To preserve confidentiality and maintain anonymity, LEAs are identified only by number in this report (LEA-1 through LEA-7). LEA-4 discontinued participation in January 2019, and LEA-7 joined the study in April

2019. Table 2.3 summarizes the characteristics of the seven participating LEAs, which include one charter, in terms of academic achievement in ELA and mathematics and select student demographics. Data in the table are from 2017–18. The table also indicates enrollment of students in the state or LEA who are in grades eligible for the CAASPP summative assessments.

Table 2.2. 2018 Eligibility Survey Invitees, Respondents, and Respondents' Interest in Study Participation

Respondent Type	Number of Invitees	Total Number of Respondents	Number of Respondents "Interested"	Number of Respondents "Possibly Interested"	Number of Respondents "Not Interested"
LEA (non-charter)	349	86 (25%)	33 (38%)	9 (10%)	44 (51%)
Charter	36	12 (33%)	8 (67%)	3 (25%)	1 (8%)
Total	385	98 (25%)	41 (42%)	12 (12%)	45 (46%)

Explanation of table contents: *Line 1 shows that we invited 349 non-charter LEAs to participate in our Eligibility Survey. Of these, 86 (or 25%) responded. Of the 86 respondents, 33 (38%) reported they would be potentially interested in participating in the Impact Case Study, 9 (10%) reported they were possibly interested, and 44 (51%) were not interested.*

Table 2.3. Characteristics of LEAs Participating in Impact Case Study

Case Study LEA #	Total Enrollment	# CAASPP Eligible Students	% Met or Exceeded ELA State Standards	% Met or Exceeded Math State Standards	% SE Dis-advant aged	% SWD	% EL
LEA-1	17,122	10,764	31%	23%	89%	8%	58%
LEA-2	4,270	2,188	59%	44%	49%	10%	26%
LEA-3	2,465	564	83%	73%	59%	9%	7%
LEA-4	4,882	2,656	30%	19%	91%	15%	37%
LEA-5	621,414	262,099	42%	32%	81%	14%	23%
LEA-6	3,926	1,976	45%	29%	47%	13%	3%
LEA-7	22,777	11,979	30%	20%	85%	12%	23%
All CA	6,220,413	3,275,552	50%	39%	60%	12%	20%

Explanation of table contents: *Line 1 shows that the LEA we labeled LEA-1 had a total enrollment (across all schools, including those not participating in the study) of 17,122 students in 2017–18. Of these, 10,764 were eligible to participate in the CAASPP summative assessments. Of those who took the summative assessment, 31% met or exceeded the ELA state standards, and 23% met or exceeded the math state standards. In LEA-1, 89% of students were socioeconomically disadvantaged, 8% were students with disabilities, and 58% were English learners.*

Readers should note that the California State Board of Education, and other states, adopted the Common Core State Standards (CCSS) in ELA and mathematics in 2010. The CCSS are more rigorous than California's previous standards and include some reorganization of content across grade levels. The Smarter Balanced Summative Assessments are aligned to the CCSS and first became operational in 2015, replacing paper and pencil assessments. Because of the substantive changes to the content standards and the time needed to implement them at the LEA and school level, the CDE anticipated the test would be very challenging to students in the initial years until adjustments to instruction caught up with the changes. In keeping with typical patterns following implementation of new standards, the statewide percentages of students meeting or exceeding the standards have been gradually increasing over time (from 2015 to 2018, an increase of 5.88% in ELA and 5.65% in mathematics) along with students' opportunity to learn the knowledge and skills measured by the assessment (CDE News Release, https://www.cde.ca.gov/nr/ne/yr18/yr18rel62.asp).

Each participating LEA signed a Memorandum of Understanding (MOU) with HumRRO agreeing to participate in a specified set of data collection activities for the duration of the 2018–19 school year. The MOU identified a point of contact (POC) for the LEA, listed the participating schools, and identified a POC for each school. The MOU stated, in summary form, the key research questions the study sought to answer, and HumRRO designed the data collection instruments to address the research questions. The MOU also stated the LEA and each school would receive a $900 honorarium for participating. HumRRO provided a portion of the full honorarium to districts that participated only part of the year (i.e., they dropped out early or joined in late). The method (e.g., gift card or check) and frequency of the honorarium (i.e., monthly or two payments) was decided by the LEA. In some cases, the LEA allowed schools to indicate their preference for how to receive the funds. One LEA turned down its honorarium.

The eligibility screening for threshold IAB usage in the sample was effective in predicting continued usage during the study year, as evidenced by information presented later in this chapter.

The DL log-on data used to screen for LEA participation was not predictive of DL usage during the study year. There were two reasons for this. First, some teachers who logged onto the system when participating in professional development activities never actually used the resources. Second, persons who entered the DL through the IA Reporting System (through "Playlist" or "Instructional Resources" links) were not captured as DL logons. HumRRO did not collect 2018–19 DL log-on data because it would be insufficient for our purposes. HumRRO will consider revising questions about use of the DL in year two of the study to capture alternate ways teachers access its resources.

Data Collection

Based on the study design, HumRRO gathered data from various sources to describe the context and use of CAASPP components by each LEA and its study schools. Though HumRRO attempted to collect all information from all participants, this was challenging because of the various levels of LEA and school participation.

HumRRO collected the following data from extant sources:

- *Statewide assessment data*. Records of summative assessment administration results and counts of interim assessments administered in each content domain.
- *Demographic records*. Data with LEA characteristics, including student population, number of schools, student demographics, and achievement on summative assessments.

HumRRO generated data about LEA and school use of CAASPP components through the following activities:

- *Data from in-person visits to LEAs and schools*. Two HumRRO researchers visited each LEA office and study school to (a) conduct interviews with the POCs and LEA and school leaders and (b) hold focus groups with teachers. Researchers took detailed notes and audio-recorded the interview and focus group responses of LEA leaders, school leaders, and teachers about the use of Smarter Balanced components of the CAASPP System. Researchers also collected artifacts such as teacher collaboration time or PLC meeting schedules, teacher lesson plans that incorporate formative tools, school calendars, handouts from student assessment data review meetings, and professional development materials. HumRRO submitted draft interview and focus group protocols (topic guides) to the CDE for review in advance of the first LEA site visit. HumRRO also provided the draft protocol to the CAASPP Technical Advisory Group (TAG) members in advance of the September 2018 TAG meeting for review and feedback. HumRRO conducted six site visits in October and November 2018 and January 2019. Because LEA-7 did not join the study until April 2019, HumRRO conducted interviews telephonically and did not conduct focus groups with teachers in that LEA. See Appendix A for an example of the interview and focus group protocols.
- *Data from monthly phone or email polling of LEA and school POCs.* For each of five months of the study, HumRRO emailed LEA and school POCs a template that included one to three questions related to the use of Smarter Balanced components as well as guidance for collecting responses. HumRRO informed the CDE of the questions asked and provided an opportunity each month for the CDE to suggest additional questions. POCs were allowed about one month to collect and submit to HumRRO narrative responses from LEA leaders, school leaders, and teachers. Due to the rolling start of cases in the study, some LEAs and schools received different questions in a particular month than did other LEAs and schools. The two late-starting cases (LEA-6 and LEA-7) did not receive the full set of monthly polling questions. See Appendix B for the full roster of school-level and LEA-level questions asked during the 2018–19 school year.
- *Data from end of school year Web-based focus groups with LEA and school POCs*. One HumRRO researcher facilitated five online focus groups: one with LEA POCs, two with elementary school POCs, and one each with middle/junior high school and high school POCs. HumRRO provided the focus group questions to participants

prior to the meeting. A second researcher took detailed notes of LEA and school POCs' responses. The focus groups were audio-recorded.

- *Data from student focus groups led by school POCs*. HumRRO asked each school POC, except those from two LEAs that did not fully participate (LEA-4 and LEA-7), to conduct one focus group with a sample of students in their school to obtain their perspectives on the various aspects of the CAASPP System. HumRRO prepared all materials for the student focus groups, including (a) parent/guardian notification letter (English and Spanish versions) that described the goal of the student focus groups and offered parents/guardians the opportunity for their students to opt out of this activity; (b) guidelines for student selection; (c) series of step-by-step instructions for conducting the student focus groups; (d) focus group script, including questions to be asked orally; (d) template for collecting student oral responses; and (e) handout of multiple-choice questions for students to answer in writing. HumRRO emailed materials to LEA and school POCs on March 22, 2019, with a request to conduct the focus groups and return student responses before the end of the school year. HumRRO received summaries of student responses from all schools for LEA-1, LEA-5, LEA-6; LEA-2 provided a summary for only one school. HumRRO received written responses to student questions from all three schools from LEA-1 and the elementary school from LEA-6. LEA-3 did not provide student focus group data due to competing priorities for staff time, LEA-4 dropped out of the study prior to the student focus groups, and LEA-7 joined the study late and was not asked to provide these data.

Data Analysis Methods

The Impact Case Study primarily involved collecting qualitative data through site visits, monthly POC polling, virtual end-of-year POC focus groups, in-person school POC-led student focus groups, and students' written questionnaire responses. HumRRO reviewed the data collected on an ongoing basis to inform questions asked during monthly polling and end-of-year focus groups. Prior to analyzing the qualitative data, HumRRO conducted several quality checks. First, immediately following each data collection activity (e.g., in-person or virtual interviews and focus groups), HumRRO researchers reviewed their notes against the audio-recording to verify accuracy of the contents and fill in any gaps. HumRRO produced Word documents of the transcribed data. Second, because LEA and school POCs provided data to HumRRO in a variety of formats and used a variety of naming conventions for their files, HumRRO reorganized and renamed data files as needed to ensure appropriate identification of the source LEA and school. Additionally, researchers cleaned the polling data files, verified responses aligned with the questions, and compiled all responses to each month's polling questions into LEA- and school-level Excel files. After the quality assurance steps were completed, HumRRO analyzed all data sources concurrently and triangulated information to describe each LEA and its schools.

HumRRO used the text analysis features of the MAXQDA software package to analyze the qualitative data collected for the Impact Case Study. First, HumRRO created and applied a naming convention to identify the LEA and school associated with each

source document. HumRRO then organized source document by file type (e.g., LEA POC interview transcripts, teacher focus group transcripts, January monthly polling responses) and formatted them to facilitate importing. Next, HumRRO researchers imported the cleaned data files into MAXQDA. The Impact Case Study director conducted an initial review of the data in each document to (a) identify major themes and (b) develop an initial list of data codes. For example, IAB Decision was identified as a main theme that included four codes: Justifications, Individual Teachers Decide, LEA Mandate or Guidance on IAB Usage, and Teacher Groups Decide. The full set of codes were reviewed and refined in an iterative fashion. The final coding system was incorporated into a single Excel document that included descriptions, and then imported into MAXQDA. HumRRO analysts used the coding system to mark text segments with similar content. Organizing and structuring the data gathered throughout the year allowed HumRRO to identify key content used to develop major themes regarding case study findings.

Five analysts were individually assigned to lead the data analysis for one or more of the seven LEAs. Each analyst began with the same MAXQDA template file, preloaded with all source documents and the coding system. Using the template file, each analyst reviewed and coded data relevant only to their LEA. Analysts reviewed all text for their LEA and its schools. If text relevant to the research questions was identified but did not fit the existing codes, analysts recommended new codes, which were shared with the other analysts so they could update their template files. The analysts communicated regularly about the coding process, especially to discuss the application of codes when the data were unclear.

For consistency in reporting the findings by LEA, the study director provided analysts a report template, along with guidance on where and how to address coded themes. Following the coding process, each analyst retrieved and reviewed coded segments to develop a draft summary of findings for their LEA(s). Two HumRRO researchers with first-hand involvement in collecting the data reviewed the LEA findings for accuracy, clarity, and consistency across sections. Analysts then reviewed, revised, and finalized their LEA sections.

HumRRO's qualitative analysis process ensured data were systematically analyzed in a manner that captured all key information shared by LEAs and schools and treated information as similarly as possible across all LEAs. Each LEA's findings follow the major themes of the research questions (contextual factors, use of summative and interim assessments, and use of the Digital Library). These detailed findings also include unique aspects about how each entity used the CAASPP System. The detailed LEA-specific findings are presented in Appendix C.

HumRRO's next step was to develop a summary for each LEA, consolidating the detailed LEA-specific findings and concisely reporting on the contextual factors, use of summative and interim assessments, and use of the Digital Library. The summaries of LEA-specific findings are presented in Appendix D.

The final analysis step involved developing summaries of major themes across all schools and LEAs. This was accomplished by reviewing each of the individual LEA-level summaries and noting common themes across the group of LEAs for each CAASPP component (i.e., summative assessments, IAs, and DL).

Overall Findings and Conclusions of the Impact Case Study

This section summarizes the experiences of collaborating LEAs and schools as evidence to respond to the 13 Impact Case Study research questions. HumRRO concludes this section with a list of best practices for using the CAASPP components and recommendations based on findings from the seven LEAs studied.

HumRRO acknowledges several limitations of the study's findings. Although the findings offer valuable insights into a very small sample of diverse LEAs and schools, these findings are not representative of LEAs and schools statewide. The study sought to identify "best" practices rather than documenting "typical" practices. The qualitative data collected reflect perceptions of study participants, which are by nature subjective. Responses to monthly polling questions at the school level were submitted by varying numbers of teachers, who were selected or recruited by the POCs, and the quantity of responses was inconsistent across schools. Although collaborating LEAs agreed to participate fully in all data collection activities, some POCs and teachers were overwhelmed with engaging in the research study and keeping up with district and school responsibilities; thus, their contributions were less than robust. Similarly, not all school POCs conducted the student focus groups. In addition, some of the responses represent limited understandings or awareness about the capabilities of the CAASPP system. We also note that some of the concerns expressed by teachers are those the CDE has already begun to address. Despite these limitations, LEA and school staff provided a wealth of firsthand accounts of their experiences working with CAASPP components during the 2018–19 school year.

School/LEA Context and Use of Full Suite of CAASPP Components

According to the primary theory of action for the CAASPP program, the Smarter Balanced components—working together to accurately assess student achievement relative to grade level curriculum standards (i.e., the CCSS)—give information to educators to help improve instruction and thus improve student achievement. The Impact Case Study examined LEAs who are implementing the full system of components to explore how the theory of action for CAASPP components might be at work and driving efforts for improving student achievement. The theory states that educators who use information from the system of components support high expectations, increase learning opportunities for students, and take advantage of curriculum and instructional materials and rich professional development resources to help effectively teach the content embodied by the standards.

1. **What are the characteristics and contexts of sampled schools/LEAs that have implemented the full suite of Smarter Balanced components?**

Although the plan was to identify a demographically diverse set of LEAs to participate in this study, the ultimate focus was to identify strong, collaborative CAASPP implementers. Some geographic diversity was achieved, with four of the selected LEAs located in southern California, two in central California, and one in the northeastern part of California. The LEA sample includes a direct-funded charter school, to represent the charter perspective. The sample also offers some diversity in student achievement, with two districts performing better overall on the ELA and mathematics Smarter Balanced summative assessments compared to the state average, and five districts performing below the state average. The LEAs reported three percent (LEA-6) to almost 60 percent (LEA-1) EL students. The majority of students in five LEAs qualified for free or reduced lunch. The LEAs reported similar levels of students with disabilities to each other and to the state average.

One common characteristic across participating schools was strong support for CAASPP component use at the LEA level. While this often extended to the school-leader level, there were mixed perspectives at the teacher level. LEA- and school-leader staff agreed to participate in the study because they believed there was value to what the CAASPP System had to offer. They were happy to support research they felt could benefit their LEA and schools as they move forward with CAASPP implementation, as well as provide feedback to improve the CAASPP System.

HumRRO noted strong teacher collaboration cultures (e.g., active PLCs) across most of the LEAs and schools, with many schools allotting time for teachers to meet regularly in grade-level or subject-area groups. These meetings generally involved instructional planning and sometimes included identifying and scheduling interim assessments and discussing assessment results. Collaboration among teachers in some cases had a positive effect for encouraging and increasing CAASPP component use. For example, two elementary school teachers in LEA-6 were strong supporters of CAASPP component use, including use of IABs that required hand scoring and use of the DL Playlist based on student IAB performance. Though not all teachers at the school embraced CAASPP to the same extent, these teachers' enthusiasm was effective in encouraging other teachers to use IABs and understand their benefits. Similarly, teachers at LEA-1 schools made ultimate IAB decisions together in PLC settings and worked together to identify instructional goals and classroom assessments. CAASPP components at these schools were used consistently at a grade level. In schools where there was not sufficient time to support PLC work, there was less opportunity to share strategies for using the CAASPP components. For example, teachers at a high school in LEA-6 did not have release time for PLC meetings. This school was one for which the study POC found value with IABs and used them; however, other teachers in the school did not.

Student access to technology (i.e., laptops or tablets with keyboards, online curricular components) varied across the LEAs. Most LEAs had enough Chromebooks in the classroom for each student, and other LEAs supplemented the number of classroom

units with shared computer carts. However, middle and high schools in LEA-5 and schools in LEA-6 did not have a one-to-one ratio of students to laptops.

2. **How does implementation of Smarter Balanced components vary across schools/LEAs? What instructions and supports are provided to educators for implementing the components?**

There were some consistencies in how Smarter Balanced components were implemented across all participating LEAs. One characteristic of implementation was the degree of support LEA-level leadership provided. Leadership personnel at all participating LEAs were supportive of the CAASPP System and promoted strong use of its components at the school level. Particularly, LEAs provided data for schools' review of and reflection on summative assessment results, and LEAs encouraged or required schools to administer interim assessments.

Another commonality across LEAs was the extremely limited use of DL resources via direct log on to the DL. HumRRO did not study whether educators accessed the DL indirectly through the Instructional Resources button in the IA Reporting System. Most teachers across schools and LEAs reported not using DL resources. There were some exceptions. One teacher at the LEA-6 elementary school used resources identified through the DL Playlist, and one LEA-3 middle school teacher noted use of materials for Algebra. Though most teachers in the study indicated familiarity with the DL, and many had logged on at one time or another, teachers had trouble finding the time to explore the DL and incorporate the resources into their already full curriculum and other available resources.

Though there were key similarities regarding professional development opportunities, there were also important variations. For example, at most schools the CAASPP coordinator and sometimes administrators had received formal training on use of CAASPP components offered by the CDE. An exception was LEA–1, which sent a small number of teachers to the CAASPP Summer Institute. The LEAs did not have funding to allow all, or in some districts any, teachers to attend formal training, but in many schools their CAASPP coordinators provided professional development locally. One school in LEA-5 indicated they have had trouble finding time to provide educator-level training on CAASPP component use, and only passed along information informally during conversations or through email communication. Another district (LEA-7) provided CAASPP training to all teachers at the beginning of the year through their school CAASPP coordinators.

The researchers found great variations in the implementation of IABs within and between schools and LEAs in our study. HumRRO attributes much of the variation to different decision-making policies about IABs. Though it was commonly felt that IABs were a positive aspect of the CAASPP System, the selection of specific IABs to administer and when to do so depended to a great extent on guidance or mandates by LEA and school leadership. For example, LEA-1 allowed decisions to be made by grade-level PLCs. Some grade levels chose to use many IABs, and others chose to use few or none. In LEA-2, IABs were also optional, and the high school and middle school

chose not to use any IABs in 2018–19, though the elementary school embraced them. The schools of LEA-7 indicated using many IABs throughout the school year. Teachers in LEA-5 and LEA-4 were required to administer a small number of IABs per subject per year; some teachers chose to administer more, while others stuck with the minimum. However, LEA-5 mandated the particular IABs, while LEA-4 allowed teachers to select their own. There was also variation in the manner of administration – some only administered IABs in a standardized manner, others in nonstandardized manner, and still others used a combination of both administration types, each for different purposes. Nonstandardized manner of administration includes use of the IABs for formative assessment or for instructional purposes.

3. **What aspects of Smarter Balanced components are perceived as most beneficial for improving classroom instruction and student learning across schools/LEAs?**

Most commonly, teachers and POCs across LEAs referred to the interim assessments, particularly the IABs, as the most useful and beneficial CAASPP component. In general, teachers across schools indicated IABs were useful for familiarizing students with the rigor and format of the summative assessment. Some teachers also noted benefits from using results to inform instructional decisions. For example, teachers at one elementary school in LEA-1 used results information to identify gaps in student knowledge and to target instruction. And teachers at LEA-3 used IAB results as evidence of how well their curriculum aligned to the summative assessment and to identify gaps in knowledge that additional instructional resources could address. Teachers also indicated that giving IABs in a nonstandardized manner (e.g., allowing students to work through the assessments in groups or as a full class) was useful as a guided instructional activity.

Perspectives on the utility of summative assessment results for improving classroom instruction and student learning varied. Most school leadership indicated summative assessments were important for generating annual goals for improving achievement levels. The LEA-5 elementary school required teachers to use summative assessment results to create grade-specific action plans. At the teacher-level, many teachers across the schools indicated the results came too late to be useful. However, teachers in multiple schools—including the elementary school teachers from LEA-2 and LEA-6—were influenced to increase the rigor of classroom instruction because they knew continuing only with their main curriculum would not be adequate to prepare their students for the content demands of the summative assessment. For example, one LEA-2 elementary teacher responded to summative results by exposing students to assignments that required incorporating information from multiple texts in informational writing. Similarly, teachers across schools and LEAs regularly cited their use of IABs was driven by their hope to improve summative assessment scores.

4. **What changes to the Smarter Balanced components and supporting resources do LEA and school staff believe would improve support for their use of CAASPP components to improve classroom instruction and student learning?**

Participants made several recommendations for improving the utility of CAASPP components. For example, many teachers and school POCs generally liked the interim assessments but felt their use could be improved by (a) increasing the number of available IABs (i.e., alternate forms of the same IABs, new IABs that address skills currently offered only in IABs that require hand scoring) and (b) removing the burden of hand scoring. Many teachers had limited use of the DL; feedback included struggles with logging in and lack of time to sort through and find useful information. One teacher had searched the DL but did not find sufficient high school ELA materials. Thus, increasing the number of resources at the high school level may increase usage of the DL by at least some teachers.

Multiple teachers indicated they would benefit from additional professional development regarding CAASPP components. Some teachers requested changes or additional features of components or resources (e.g., item level analysis for IABs, teacher access to a classroom's IAB results), unaware they already exist. For example, many teachers and school leaders were unfamiliar with enhancements made to the IA Reporting System since the initial launch of IABs in 2015. This lack of knowledge may be a result of lack of training or because at those schools, IAB results were uploaded to an LEA-wide information system that included other assessment and student data. Some teachers expressed an interest in learning how IABs could be used in nonstandardized ways for instructional purposes. Though the CDE offers various in-person professional development training opportunities, oftentimes an LEA lacks funding to send teachers. Or, the LEA lacks time to convey the learning experiences or review handouts, presentation, and tools with teachers. Teachers may benefit from additional efforts by CDE, their LEA, and/or school leadership to ensure valuable information, available in many online CAASPP resources, is communicated to them for implementation at the classroom level.

Technology issues were also discussed as a challenge that could be improved. For example, teachers would prefer to access all CAASPP related information through one sign-on. Also, CAASPP coordinators noted the large amount of time required to complete rostering that allows teachers to view IABs results. Many indicated the time required for this task resulted in delays in teachers having access to scores. LEA-7 noted the text-to-speech feature pauses in odd places. Some teachers wanted the IAB hand scoring process to be less involved technologically. These teachers cited using other assessment resources with constructed response items (e.g., those with rubrics) that could be more readily and easily hand scored, with scores directly entered into their LEA's student information system.

5. **How do educators/schools/LEAs use and integrate results from the summative, interim, and formative assessment resources for each content domain (ELA/literacy and mathematics) with each other and with other measures to enhance classroom instruction and student learning? What challenges are faced and how are they overcome?**

The studied LEAs used multiple measures to evaluate student progress, in addition to CAASPP components. For example, teachers in LEA-1 used a combination of CAASPP IABs and tests they generated in SchoolCity to support their instructional needs. Other LEAs described using Illuminate tests, including tests teachers created, as well as district-wide assessments, such as NWEA. Teachers also used short formative assessments available through their curriculum at the conclusion of direct instruction on a topic. These various measures were used to monitor student progress and identify gaps in knowledge and skills. Some teachers used primarily non-CAASPP resources for these purposes and used IABs primarily to prepare for the summative assessment; however, others used information from both IABs and non-CAASPP resources to track student performance and progress.

At the high school level, CAASPP results and other measures were used to assist with student placement decisions. For example, in LEA-4, counselors used grade eight ELA results from feeder schools to help make decisions about student placement into grade nine and ten ELA support classes; however, NWEA results were used to place students into the grade nine mathematics support class. School leaders noted CAASPP Early Assessment Program (EAP) results typically arrive after decisions about student placement into grade twelve “college transition” courses have been made; however, counselors may use EAP scores to adjust enrollment in these courses during the school year.

Teachers cited several challenges with integrating assessment results to enhance classroom instruction and student learning. Specifically, some teachers indicated too much time is required to administer the IABs and requested shorter IABs, to include less administration time and fewer items. Additionally, existing IABs sometimes include more than what a teacher has taught and wants to measure, or they may include skills taught at a different time during the school year. Also, some skills are measured only in IABs that require labor-intensive hand scoring. Teachers who use IABs as pre- and post-tests (after instruction targeting the skills measured by the IAB) request multiple forms of the same IAB to effectively use the tests for this purpose. Despite our finding that this was a relatively common use of IABs, the CDE cautions about using an IAB for a pre- and post-test comparison. Specifically, there is concern that the IABs are relatively brief and teachers should consider multiple sources of information to understand student learning. In addition, teachers should consider whether the full content of an IAB will be taught between the pre- and post-test if used for this purpose.

Challenges related to time constraints and student population were also described. Teachers noted additional release time to collaborate in PLCs was necessary to allow for in-depth review of assessment results and planning appropriate actions to respond to them. Some school leaders and teachers described the challenge of targeting

instruction at the rigor of the summative assessments when students are below or far below grade level.

6. **How do students from schools that use the full suite of Smarter Balanced components perceive classroom opportunities to learn about summative assessment item types and topics for each content domain (ELA/literacy and mathematics)?**

The small number of students who participated in focus groups (conducted by school staff) after students took the spring 2019 summative assessments indicated familiarity with Smarter Balanced item types and the test format. Students indicated, in part, this was because they had prior opportunities to practice on similar computer-based assessments. Though most expressed comfort with the item types, some expressed beliefs that items were written in a tricky or confusing manner. Student responses were mixed regarding familiarity with the summative topics. Though some found the topics easy and consistent with what they learned in class, others had not seen or did not remember being taught certain topics. One student noted difficult vocabulary, another expressed they had not been taught measurement topics during the school year, and another was not prepared for the level of depth they were expected to go into in their writing.

Smarter Balanced Summative Assessment

One primary purpose of the Smarter Balanced summative assessments is to provide valid, reliable, and fair information about students' ELA/literacy and mathematics achievement with respect to CCSS, for students in grades three to eight and high school. The following research question explored how LEAs and schools used the data from the 2018 summative assessment during the 2018–19 school year.

7. **How do educators/schools/LEAs use summative assessment data—including, but not limited to, information about student proficiency levels and progress towards college- and career-readiness—in ELA/literacy and mathematics to inform classroom instruction and make decisions?**

School POCs and teachers participating in the Impact Case Study reviewed summative assessment data from the prior year during the first semester of the current school year. Some schools reviewed data as a school-wide team early in the year, while other schools did not do so until November or December. The degree to which data was reviewed and used varied among schools. Almost all school leaders and teachers at the elementary and middle schools reviewed grade-level results of the percentage of students that tested at each overall achievement level. Many also reviewed claim level results, and a few accessed target reports. Many teachers indicated receiving Student Score Reports (SSRs) from students they taught the previous year but did not find this information useful because the content areas of performance were so broad. Others noted they received scores for their current students but not scores from students they taught the previous year, so they were unable to consider scores when reflecting on their prior year's teaching. According to the CDE, the reporting system allows teachers

to see groups of students based on what access teachers are provided by their administration. Teachers may not have access to adjacent grade levels in a current year; however, teachers are able to independently download class data every year and thus examine performance over time. In addition, teachers can examine test histories for individual students or groups for whom they have access, going as far back as the 2014–15 school year.

In some cases, teachers had difficulty recalling data they reviewed, whereas others indicated summative results were instrumental for developing instructional plans and goals. Schools in LEA-7 used summative results as one piece of evidence for setting growth goals at the group (e.g., grade level) and student level. During a school-wide meeting, teachers of the LEA-5 elementary school developed action plans that were directly influenced by summative assessment scores. At the district level, LEA-5 used students' summative results to identify for each grade and content area the specific IABs that would be mandated across the LEA.

Some LEAs and schools focused their attention on how far away their students were from the overall "Standard Met" achievement level. They discussed goals in terms of what needed to be done to increase the number of students meeting or exceeding ELA and mathematics standards. LEA-5 reported a focus on the distance from meeting the standard at the school level, and some teachers shared information with individual students, illustrating their distance from meeting the standard. LEAs did not describe monitoring scores that specifically indicate college readiness; a low number of high schools participated in the study, which may have limited the likelihood of identifying uses of summative assessment results for this purpose.

Teachers from some schools had trouble recalling whether and when data were shared. At an LEA-6 elementary school, some teachers recalled receiving summative results, whereas others did not recall receiving them. Some school staff did not use the summative results because they didn't receive them in a timely manner; they often did not receive the summative results until they were well into the new school year.

Interim Assessments and Digital Library Resources

One of the Professional Learning resources in the DL is called "Understanding the Smarter Balanced Interim Assessments." This excerpt from the resource describes research supporting the value of interim assessments: "While a rigorous summative assessment is important, it is insufficient to drive all of the change in teaching and learning. Informed by experiences in England and Hong Kong, interim and formative assessments are the other necessary assessment ingredients to drive teaching and learning (Darling-Hammond and Pechone, 2010). Grounded in cognitive development theory about how learning progresses across grades and competence develops over time (NRC, 2001; Pellegrino, 2006), Smarter Balanced interim assessments: (a) work in concert with the summative assessment; (b) allow for more innovative and fine-grained measurement of student progress toward the Common Core State Standards (Shepard, Hammerness, Darling-Hammond and Rust, 2007); and (c) provide diagnostic information that can help tailor instruction and guide students in their own learning efforts."

The following research questions explored several aspects of how LEAs and schools used the interim assessments and the DL during 2018–19.

8. **What interim assessments are used for ELA/literacy and mathematics for schools/LEAs that have implemented the full CAASPP System, and at what grade levels and frequency?**

IABs were used by all study schools in 2018–19 except for one high school in LEA-2. Table 2.4 presents, for the state of California overall and for each LEA, the average number of IABs administered per school, among only the schools choosing to use IABs. Schools not using any IABs were not included in the averages. The table provides the average number of ELA and mathematics IABs administered overall, as well as the number administered by subject area and the number administered in a standardized or nonstandardized manner. Three case study LEAs greatly exceeded the state average for IAB usage (LEA-2, LEA-3, and LEA-5), LEA-1 was very close to the state average, and two LEAs were below the state average (LEA-6 and LEA-7). LEA-2 administered far more IABs in nonstandardized manner than in standardized manner, while LEA-3 and LEA-5 administered IABs in each manner about equally. Only LEA-3 did not administer any ELA IABs. As a reminder, LEA-4 discontinued participation in January 2019, and LEA-7 joined the study in April 2019.

Table 2.4. Average Number of Smarter Balanced IABs Administered Per School, Statewide and by Case Study LEA, and by Subject Matter and Manner

	# Schools Giving IABs	Average # IABs Per School ELA and Math	Average # IABs Per School ELA	Average # IABs Per School Math	Average # Standardized IABs Per School (ELA and Math)	Average # Non-Standardized IABs Per School (ELA and Math)
All California	6,211	1,334	577	757	768	566
LEA-1	19	1,379	467	912	802	576
LEA-2	3	2,731	1,576	1,155	419	2,312
LEA-3	1	2,407	0	2,407	1,031	1,376
LEA-4	7	1,011	374	637	761	250
LEA-5	726	1,944	845	1,099	937	1,007
LEA-6	10	805	395	410	550	255
LEA-7	24	1,195	439	757	445	750

Explanation of table contents: *Row 1 shows across all of California there were 6,211 schools that gave IABs during the 2018–19 school year. For these 6,211 schools, the average number of total IAB administrations was 1,334. Schools giving IABs in California on average gave 577 ELA IABs and 757 math IABs. They gave 768 IABs in a standardized manner and 566 in a nonstandardized manner (across math and ELA).*

Most study schools administered ELA and mathematics IABs, although LEA-3 (charter high school) chose to administer only mathematics IABs. Appendix C includes LEA-specific tables of IAB usage information, summarizing how many times specific IABs at each grade level were administered by participating schools. Here, tables 2.5 through 2.7 summarize across the seven LEAs and their three levels of schools in the study (elementary, middle, high school) the number of times ELA IABs were offered, by test name and grade. The count of opportunities refers to the number of test administration sessions for each specific IAB (e.g., Brief Writes). This is a count of the number of times the test was offered (e.g., to one classroom, to a special subset of students), not the number of students who participated in each administration of the test. In the tables, NA indicates the IAB is not available at that grade. At the elementary school level, the Performance Task (which is unique for each grade) was the most frequently offered ELA IAB. At the middle school level, it was Read Informational Texts, and at the high school it was Research.

Table 2.5. Count of Opportunities to Take Specific IABs in English Language Arts, Across Elementary Schools in the Impact Case Study

Test Name	Grade 3 (N Schools=10)	Grade 4 (N Schools=10)	Grade 5 (N Schools=10)	Totals
Brief Writes*	11	5	10	26
Editing	7	7	11	25
Language and Vocabulary Use	12	12	11	35
Listen/Interpret	12	7	5	24
Read Informational Texts*	30	17	20	67
Read Literary Texts*	21	17	20	58
Research	5	5	5	15
Revision	3	1	1	5
Performance Task*	48	39	34	121
Totals	149	110	117	376

* Indicates IAB includes some open-ended responses that require hand scoring, if the test is administered in standardized manner.

Explanation of table contents: *These opportunities may have been a full class session or a session for a select group of students. Row 1 shows that for the schools in our study only, there were 11 opportunities (i.e., test sessions) for Brief Writes at grade 3, 5 opportunities at grade 4, and 10 opportunities at grade 5. Overall, across all our study schools, there were 26 opportunities to take Brief Writes in the elementary grades 3 through 5.*

Table 2.6. Count of Opportunities to Take Specific IABs in English Language Arts, Across Middle Schools in the Impact Case Study

Test Name	Grade 6 (N Schools=6)	Grade 7 (N Schools=6)	Grade 8 (N Schools=6)	Totals
Brief Writes*	2	3	1	6
Editing	11	5	0	16
Language and Vocabulary Use	12	9	0	21
Listen/Interpret	6	5	5	16
Read Informational Texts*	13	17	13	43
Read Literary Texts*	10	10	4	24
Research	2	8	5	15
Revision	5	4	0	9
Edit/Revise	NA	NA	4	4
Performance Task*	6	3	4	13
Totals	67	64	36	167

* Indicates IAB includes some open-ended responses that require hand scoring, if the test is administered in standardized manner.

Explanation of table contents: *These opportunities may have been a full class session or a session for a select group of students. Row 1 shows that for the schools in our study only, there were 2 opportunities (i.e., test sessions) for Brief Writes at grade 6, 3 opportunities at grade 7, and 1 opportunity at grade 8. Overall, across all our study schools, there were 6 opportunities to take Brief Writes in the middle school grades 6 through 8.*

Table 2.7. Count of Opportunities to Take Specific IABs in English Language Arts, Across High Schools in the Impact Case Study

Test Name	High School (N Schools=4)
Brief Writes*	4
Editing	4
Language and Vocabulary Use	5
Listen/Interpret	4
Read Informational Texts*	4
Read Literary Texts*	3
Research	10
Revision	8
Performance Task*	1
Totals	43

* Indicates IAB includes some open-ended responses that require hand scoring, if the test is administered in standardized manner.

Explanation of table contents: *These opportunities may have been a full class session or a session for a select group of students. Row 1 shows that for the schools in our study only, there were 4 opportunities (i.e., test sessions) for Brief Writes in high school.*

Tables 2.8 through 2.10 summarize the number of times mathematics IABs were offered, by test name and grade. At the elementary school level, Number and Operations in Base Ten was the most frequently offered mathematics IAB. At the middle school level, it was The Number System, and at the high school level it was Algebra and Functions I.

Table 2.8. Count of Opportunities to Take Specific IABs in Mathematics, Across Elementary Schools in the Impact Case Study

Test Name	Grade 3 (N Schools=10)	Grade 4 (N Schools=10)	Grade 5 (N Schools=10)	Totals
Measurement and Data	21	5	4	30
Number and Operations - Fractions	18	19	13	50
Number and Operations in Base Ten	24	29	26	79
Operations and Algebraic Thinking	34	9	14	57
Geometry	14	6	12	32
Performance Task*	12	9	11	32
Totals	123	77	80	280

* Indicates IAB includes some open-ended responses that require hand scoring, if the test is administered in standardized manner.

Explanation of table contents: *These opportunities may have been a full class session or a session for a select group of students. Row 1 shows that for the schools in our study only, there were 21 opportunities (i.e., test sessions) for Measurement and Data at grade 3, 5 opportunities in grade 4, and 4 opportunities at grade 5. Overall, across all our study schools, there were 30 opportunities to take Measurement and Data in the elementary grades 3 through 5.*

Table 2.9. Count of Opportunities to Take Specific IABs in Mathematics, Across Middle Schools in the Impact Case Study

Test Name	Grade 6 (N Schools=6)	Grade 7 (N Schools=6)	Grade 8 (N Schools=6)	Totals
Geometry	1	4	5	10
Expressions and Equations	9	18	NA	27
Ratios and Proportional Relationships	14	9	NA	23
The Number System	15	11	8	34
Statistics and Probability	1	2	NA	3
Expressions and Equations I	NA	NA	17	17
Expressions and Equations II	NA	NA	3	3
Functions	NA	NA	11	11
Performance Task*	5	5	2	12
Totals	45	49	46	140

* Indicates IAB includes some open-ended responses that require hand scoring, if the test is administered in standardized manner.

Explanation of table contents: *These opportunities may have been a full class session or a session for a select group of students. Row 1 shows that for the schools in our study only, there was 1 opportunity (i.e., test session) for Geometry at grade 6, 4 opportunities in grade 7, and 5 opportunities at grade 8. Overall, across all our study schools there were 10 opportunities to take Geometry in the elementary grades 6 through 7.*

Table 2.10. Count of Opportunities to Take Specific IABs in Mathematics, Across High Schools in the Impact Case Study

Test Name	High School (N Schools=4)
Statistics and Probability	6
Algebra and Functions I	14
Algebra and Functions II	10
Geometry and Right Triangle Trigonometry	4
Interpreting Functions	13
Number and Quantity	2
Seeing Structure in Expressions/Polynomial Expressions	2
Geometry Congruence	2
Geometry Measurement and Modeling	2
Performance Task*	1
Total	56

* Indicates IAB includes some open-ended responses that require hand scoring, if the test is administered in standardized manner.

Explanation of table contents: *These opportunities may have been a full class session or a session for a select group of students. Row 1 shows that for the schools in our study only, there were 6 opportunities (i.e., test sessions) for Statistics and Probability in high school.*

The statewide usage of ICAs (including only California schools administering at least one ICA) was far lower than that for IAB usage. Table 2.11 provides information for the administration of ICAs during 2018–19. LEA-4, LEA-5, and LEA-6 administered a substantive number of ICAs, with the large majority of ICAs in LEA-4 and LEA-6 administered in a standardized manner. LEA-4 administered more than twice the state average of ICAs per school. LEA-2 and LEA-3 did not administer any ICAs. As a reminder, LEA-4 discontinued participation in January 2019, and LEA-7 joined the study in April 2019.

Table 2.11. Average Number of Smarter Balanced ICAs Administered Per School, Statewide and by Case Study LEA

	# Schools Giving ICAs	Average # Total ICAs Per School ELA and Math	Average # ICAs Per School ELA	Average # ICAs Per School Math	Average # Standardized ICAs Per School (ELA and Math)	Average # Non-Standardized ICAs Per School (ELA and Math)
All California	1,262	221	99	122	169	52
LEA-1	1	1	0	1	0	1
LEA-2	0	0	0	0	0	0
LEA-3	0	0	0	0	0	0
LEA-4	3	380	183	197	349	31
LEA-5	161	157	60	97	88	69
LEA-6	2	70	11	60	60	10
LEA-7	4	2	1	1	1	2

Explanation of table contents: *Row 1 shows that across all of California 1,262 schools gave ICAs during the 2018–19 school year. For these 1,262 schools, the average number of total ICA administrations was 221. Schools giving ICAs in California on average gave 99 ELA ICAs and 122 math ICAs. They gave 169 ICAs in a standardized manner and 52 in nonstandardized manner (across math and ELA).*

9. **What decision-making processes are used by educators/schools/LEAs to determine what ELA/literacy and mathematics interim assessments to use, who should administer them, and how frequently they should be administered?**

Impact Case Study LEAs took different approaches in the decision-making process to determine IAB use. LEA-5 mandated two specific ELA and two specific mathematics IABs be administered per grade in all its schools during 2018–19; however, schools could submit a waiver request asking to substitute an alternate IAB. LEA-7 mandated each student take one ELA and one mathematics IAB during the school year. LEA-4 mandated elementary grades three through five administer the ELA and mathematics ICAs, mandated middle school grades administer at least one IAB, and mandated grade eleven administer two IABs. LEA-6 did not have a mandate but expected teachers to help all students encounter at least two IABs prior to taking the summative assessments. All LEAs with mandates also allowed schools and teachers to administer more IABs. LEAs 1, 2, and 3 did not mandate IAB use, rather they strongly encouraged them. For IAB and ICA decisions made at the school or teacher level, decisions were typically made in a PLC setting with all grades within a school administering the same IABs, though this was not always the case. Some schools included only one teacher per

grade, and other schools indicated decisions were made at the teacher level. LEAs that chose not to mandate IA use indicated they wanted to give the control to their teachers, empowering them to choose the most effective ways to incorporate IAs in the classroom. These LEAs believed success was more likely if schools and teachers drove the decisions. However, one LEA-7 school POC stated staff appreciated a recommended IAB schedule provided by school leadership.

Teachers described three key factors that were considered in deciding which IABs to use and when: (a) scope and sequence of instruction and how the skills taught align with the knowledge and skills measured by an IAB, (b) coordinating timing of IAB administration to allow adequate time to address identified student weaknesses prior to administering the summative assessment, and (c) whether hand scoring is involved. While the latter is considered an advantage by some teachers, it is considered a drawback by others. When decisions were made by a PLC process, IABs were sometimes chosen based on a review of summative results, to address areas identified as deficiencies among student groups (e.g., grade level).

Classroom teachers in the study schools administered ELA and mathematics IABs to their students. Specifically, ELA and mathematics teachers for grades three through eight administered IABs. At only one school (the LEA-6 elementary school), did teachers from content areas other than ELA and mathematics, including science and history, administer IABs.

10. **To what extent have educators/schools/LEAs incorporated ELA/literacy and mathematics IABs into their classes? What, if any, classroom assessments have been replaced in the process? Why, and what are the implications?**

As indicated in Research Question 8 above, most schools in the Impact Case Study administered many IABs. Most schools indicated the IABs have replaced at least some classroom assessments in recent years. However, most schools also used formative assessments found through the curriculum, and some schools combined use of IABs with other computer-based interim assessments they have access to, or those they develop themselves through programs such as SchoolCity or Illuminate. LEA-1 uses CAASPP IABs and locally developed interim assessments throughout the district. The individual PLCs determine the frequency of each, and in some cases grade-level groups have decided to replace most interim assessments with IABs. LEA-7 uses Illuminate to compile assessment data from various sources; therefore, teachers can examine student performance on CAASPP IABs and district screeners together to make decisions.

Some teachers preferred IABs over other available classroom assessments because they more closely align with what students are expected to do on the summative assessment. Multiple teachers and school leaders indicated the exposure to the rigor and format is beneficial to students.

11. **How do educators/schools/LEAs use information from ELA/literacy and mathematics interim assessments to track individual student progress and/or inform classroom instruction?**

Schools in the Impact Case Study varied in their use of interim assessments to track student progress and inform classroom instruction. Some schools indicated using IABs only (a) because they were mandated or (b) to prepare their students for the rigor, content, and/or platform of the summative assessments. For many schools in which IABs were used only for these purposes, the LEA intended to eventually increase their use as an instructional tool. However, multiple teachers had used interim assessment content and results to inform classroom instruction. LEA-4 teachers noted IABs provided a quick return of results, making them useful to identify skills to reteach. LEA-5 noted using IABs for similar purposes. Some PLCs in the district elected to administer the same IAB twice as a pre-test/post-test, and then reteach skills that students still struggled with following the second administration. LEA-1 used information learned through IAB reports to inform development of distributive practice activities (i.e., daily, short instruction sessions). Others used the interim assessment items as guidance for the level of rigor they need to incorporate into the classroom.

12. **How is information on student/school/LEA performance on ELA/literacy and mathematics interim assessments used at the school/LEA level to determine the effectiveness of practices and curricular materials for teaching the targeted standards (i.e., CCSS)?**

Though some study schools began using interim assessments to inform classroom instruction, few directly used interim assessment data to determine the effectiveness of teaching practices or curricular material. LEA-3 used IAB results as evidence of how well their curriculum aligned to the summative assessment. By examining items their students commonly missed, they were able to identify content that was not adequately covered by the curriculum. Some teachers noticed the difference in rigor between their curriculum and the summative and interim assessments. Thus, they recognized the importance of incorporating interim assessments and classroom assignments that require higher level thinking, based on Smarter Balanced assessment content.

13. **How is the Smarter Balanced Digital Library of formative tools used to improve classroom instruction (e.g., share information with students to help them monitor their own performance; better align instruction, curricula, and assessments)?**

The study schools reported extremely limited use of Digital Library resources. Many teachers were aware of the resources and had logged into the DL at least once; however, they did not have time to go through the materials or they already had sufficient materials through their curriculum. One group of teachers at an LEA-2 elementary school was not certain what the DL was, and several study participants reported confusion about and difficulty logging into the DL. In contrast, one LEA-6 elementary teacher embraced the DL. She appreciated that the links provided through

the IA Reporting System targeted her students' specific needs, based on their IAB performance.

Though the DL was cited by only two teachers as beneficial for classroom instruction, it is worthwhile to note that the schools in the study acknowledge they are only just beginning to explore use of this Smarter Balanced component. With more time to implement use of this component, including expanded teacher training and continued improvements to the DL, it might be expected these resources will increasingly be used.

Best Practices

Based on the full scope of first-year findings across the studied LEAs, HumRRO identified a sample of best practices supporting effective use of CAASPP components to improve teaching and learning. For this report, HumRRO defined a "best practice" as an approach used by participating LEAs, schools, or teachers that (a) aligns well with the intended purpose of and guidance for implementing components within the CAASPP System and (b) resulted in educators having a positive experience using the CAASPP System to inform their teaching. We believe these approaches may benefit other schools or LEAs that implement CAASPP.

- Provide support and training at the school and LEA levels for using CAASPP resources. Teachers and staff who attended CAASPP professional development or reviewed resources available online increased their comfort level with the CAASPP components, including hand scoring of IABs and using and interpreting assessment results.
- Provide leadership guidance and encouragement for using CAASPP components while allowing grade-level or content-area professional learning communities (PLCs) flexibility regarding what interim assessments and DL resources to incorporate into their classrooms.
- Facilitate school-wide data discussions to ensure teachers know how to access and interpret results, and how these data can inform instructional practices.
- Provide time and resources to support collaboration among grade-level and/or content-area PLCs to plan instruction and use formative assessments effectively.

Recommendations

HumRRO reviewed the full scope of study findings based on the perspective of the participants—a small number of teachers within a small number of schools in a small number of LEAs—to develop suggestions for the CDE to consider as part of its continuous improvement of the CAASPP System. Some suggestions are already being addressed by planned changes and updates to the system, which we describe in the section following our recommendations.

Based on the first-year findings across the case study LEAs, we offer the following recommendations:

Recommendation 1: The CDE should continue providing regional training opportunities and updated online resources for LEA- and school-level staff. The in-person trainings and CAASPP.org and CDE website resources are critical to helping educators throughout the state (a) accurately interpret Smarter Balanced summative and interim assessment results, (b) implement existing and new Smarter Balanced components, and (c) learn about enhancements to existing components.

Recommendation 2: Regarding interim assessments, the CDE should work with the Smarter Balanced Assessment Consortium to provide an expanded pool of ELA and mathematics tests, including multiple versions of existing IABs, ICAs for grades nine and ten, and shorter interim assessments that examine student achievement at the target level. Teachers using the existing interim assessments find them of high quality and requested more options for tests for classroom use.

Recommendation 3: Regarding the hand scoring requirements of some interim assessments, the CDE should explore how to address concerns related to the challenges some LEAs and schools have finding time for training and hand scoring. Some teachers in our sample who participated in hand scoring found it an excellent professional development activity, and others found instructional value in reviewing scored responses. However, constraints on time and resources often caused schools to decide against giving IABs that involve hand scoring. Perhaps the CDE could include an option for scoring via artificial intelligence techniques (currently in progress by Educational Testing Service, ETS). At the local level, support could take the form of (a) increasing the number of in-person hand-scoring training opportunities, (b) expanding the number of participants in such training, (c) providing teacher release time to engage in hand-scoring activities, or (d) sharing examples of teachers enthusiastic about their experiences with hand scoring (e.g., the CAASPP in Action series).

Recommendation 4: The CDE should encourage LEA and school leaders to provide local training opportunities, including time and resources, to help teachers (a) accurately interpret Smarter Balanced summative and interim assessment results, (b) implement existing and new Smarter Balanced components, and (c) learn about enhancements to Smarter Balanced components. LEA and school leadership receive CAASPP training on Smarter Balanced components, and sometimes these trainings are made available to teachers. However, many schools have not had the time to pass along information to all their staff. Some teachers had not tried logging on to the IA Reporting System since the many enhancements to it were launched in 2018–19. Most teachers in the small study sample had not explored the DL, often because they found logging on confusing or because they felt they had sufficient resources already. Teachers in the study who tried the DL noted frequently that navigation was difficult and time

consuming, though some of these teachers may have been referring to earlier versions of the system before it was enhanced.

Recommendation 5: The CDE should seek ways to streamline or provide additional guidance on rostering within the IA Reporting System, including recommendations regarding what access LEAs should be providing to their teachers. Some CAASPP coordinators found the CAASPP rostering process to be cumbersome, and for one LEA there was confusion in 2017–18 that resulted in teachers not having student-level results. In addition, some teachers would like more access than they are currently provided by their school or LEA. Accessibility of IA report features at the educator level is dependent upon the creation of rosters by the coordinator. Teachers may benefit if their CAASPP coordinators are given more direction regarding what level of access they should provide their teachers.

Planned CAASPP System Improvements

Several important CAASPP System improvements are planned to be implemented by the CDE in the 2019–20 school year. The following enhancements and professional development opportunities address areas of need described by LEA and school staff or observed by HumRRO:

- The CDE will host a statewide 2019 California Assessment Conference in October 2019. The three-day conference will offer a variety of sessions for classroom educators to explore the connection between assessments and classroom instruction and to explore ways of using assessment resources for improved teaching and learning.

- Beginning September 3, 2019, educators will use a single username and password (i.e., single sign on) to access the various CAASPP and ELPAC online systems, including the Test Administrator Interface, Interim Assessment Systems (Viewing System, Hand Scoring System, and Reporting System), Online Reporting System (ORS), DL, and Practice and Training Tests.

- New Smarter Balanced ICAs in ELA and mathematics will be available for administration to students in grades nine and ten in 2019–20, with different cut points for each grade level.

- New Smarter Balanced Focused IABs will measure one to three targets compared to up to eight targets measured by the current IABs. The focused IABs will measure smaller bundles of content to (a) give teachers a better understanding of students' knowledge and academic performance and (b) provide teachers with precise next steps for instruction. In addition to the more than 100 IABs already available to teachers, approximately 40 focused IABs are slated for release in 2019–20, followed by approximately 90 more over the following two school years.

- The CDE will release a new interface to the DL, currently referred to as the DL 2.0. The updated DL will address many of the concerns with the current DL. It is

expected to be easy to use, will include step-by-step directions, and will be accessible (WCAG 2.1AA compliant). The DL 2.0 is being purposefully developed to align with Smarter Balanced grade-level claims and targets and provides options and ideas for differentiation and student access of content. Embedded teachers will find formative assessment process strategies. In addition, the DL 2.0 is aligned with new Smarter Balanced quality control criteria. Finally, the DL 2.0 resources will be specifically tied to Connections Playlists, tools that link interim assessment results with teaching resources in the Digital Library to help optimize student learning.

Next Steps

The design of the Impact Case Study, as presented in the 2018 Independent Evaluation Report, includes a second year of study. HumRRO and the CDE met in spring of 2019 to review and update the study plan. Both parties agreed the second year should continue to focus on the Smarter Balanced components of the CAASPP System, because the new science assessments are still in the early stages of implementation. For the most part, the same data collection activities will be conducted, although with new LEAs and schools. The CDE had requested HumRRO seek continued collaboration with half of the LEAs (and their same schools) who participated during the first year, to provide for longitudinal study. Unfortunately, the LEAs invited to join again were unable to recommit the time and energy to another year of study. HumRRO is therefore following the model described in the study design to recruit six new LEAs for the 2019–20 school year. The selection process will include reviewing responses to an updated 2019 Eligibility Survey, reviewing IAB usage data from 2018–19, and conducting discussions with interested LEA leaders.

This page is intentionally blank.

Chapter 3: California Science Test Alignment Study Update

HumRRO approaches alignment studies as one means to gather evidence to demonstrate the validity of intended interpretations and uses of the assessment scores. Alignment studies can tell us how well a set of test items fully samples the construct represented by the associated content standards. That is, alignment studies indicate whether a test effectively measures what it is intended to measure.

For the California Science Test (CAST), evaluating alignment represents a significant challenge because of the nature of the content and the content standards. The California Next Generation Science Standards (CA NGSS) provide a framework for science education. Within the CA NGSS, performance expectations (PEs) are assessable statements of what students should know and be able to do. The following three major components, also referred to as dimensions, are combined to operationalize the PEs:

1. Disciplinary Core Ideas (DCIs) are the key ideas in science that have broad importance within or across multiple science or engineering disciplines. These core ideas build on each other as students progress through grade levels. The DCIs are grouped into the following domains: Physical Sciences, Life Sciences, Earth and Space Sciences, and Engineering, Technology, and the Application of Science (hereafter, Engineering).

2. Crosscutting Concepts (CCCs) help students explore connections across the four domains of science mentioned above in item 1. When these concepts, such as “cause and effect,” are made explicit for students, they can help students develop a coherent and scientifically based view of the world around them.

3. Science and Engineering Practices (SEPs) describe what scientists do to investigate the natural world and what engineers do to design and build systems. The practices better explain and extend what is meant by “inquiry” in science and the range of cognitive, social, and physical practices that it requires. Students engage in practices to build, deepen, and apply their knowledge of core ideas and crosscutting concepts.

The three major components of the CA NGSS (DCIs, CCCs, and SEPs) are integrated into the three science disciplines (physical sciences, life sciences, and earth and space sciences). In the CAST test design, each of the three science disciplines (physical sciences, life sciences, and earth and space sciences) assesses engineering, technology, and application of science. The design of the test is further complicated by the premise that students’ knowledge is expected to be integrated and to accumulate to create a deep understanding of science content. Students are expected to apply their knowledge and generalize across the three major components. Developing tests and test items that adequately sample such complex and integrated content is especially challenging. When an item measures a single standard or concept, the alignment process is relatively straightforward. However, test development and alignment become more complex when standards are designed as interactions among statements about content.

The first step in conducting the investigation for CAST alignment was to investigate the nature of the assessment itself, how the standards guided the development of the test items (and how the standards and items should therefore relate to one another), and the interpretations to be made from CAST scores. HumRRO then modified traditional alignment methods to account for the test structure and design, a process in keeping with best practices in test validation that facilitates using alignment study results in an overall validity argument. This process also supports federal peer review goals.

It should be noted that the operational test design for 2018–19 does not represent the final design. Future changes to the test design will be presented to the SBE for consideration. For example, the first segment (Segment A1) of the assessment may be used as a screener in the future. Students would then be administered a second part of the assessment (Segment A2) based on the student's ability level, as estimated from Segment A1. For 2018–19, all students were administered the same Segment A1 and A2 to mirror the format, but without an operational screener.

The CAST is a computer-based, fixed-form (non-adaptive) assessment administered to students in grades five, eight, and once in high school (i.e., grades 10, 11, or 12). The CAST was field-tested in spring 2018 and administered operationally for the first time in January–July of 2019. The 2019 assessment included the following three segments:

- *Segment A:* a set of selected response and short constructed-response items (two segments (A1 and A2) were administered operationally in 2019).

- *Segment B:* a set of two performance tasks (five performance tasks were available for Segment B in grade five, six in grade eight, and three in high school; two were selected per test form.).

- *Segment C:* a set of items comparable to Segment A or B, highly matrixed across test forms, each taken by a smaller sample of students than Segments A or B. Segment C included only field test items (discrete and PT), not operational items.

For the 2019 CAST, results from the first two segments were used to report individual student scores. Segment C was not used for individual score reporting but for field test purposes only. The high level test design planned for a portion of Segment C to include operational items that would provide school- and LEA-level information about student achievement on a broader sample of content than would be possible otherwise. At the time of this alignment investigation, only Segments A and B were administered operationally. All results will be based on Segments A and B only.

Because students who took the test in 2019 could potentially have been administered any combination of Segments A and B, a student testing event (or test form) was defined in this evaluation as any possible combination of Segments A and B. This means that there were 10 possible forms for grade five, 12 for grade eight, and 6 for high school. Alignment analyses will be conducted for each potential form and the results will be summarized across forms.

The structure of CAST required that HumRRO consider alignment in two ways. First, individual students' scores should be sufficiently valid and reliable to support their intended interpretations, and this could only happen if CAST adequately sampled the CA NGSS in the first two test segments (segments used to compute student-level scores). One level of alignment reporting will therefore be based only on Segments A and B. For the remainder of this report, this will be termed "student-level alignment." Analyses will directly address only student-level alignment.

The high-level test design indicates CAST results will also be reported at the group level (e.g., results by LEA) in the future. The CDE and ETS will collaborate on how to derive and report group scores (e.g., adding Segment C items to Segments A and B) using test data to inform those decisions. HumRRO expects test reporting to be less specific at the student level than at the group level. Inferences made at the group level may reference a larger set of operationally administered test items, including those in Segment C. Future alignment reporting should take the full operational item pool into account. For the remainder of this report, this will be termed "overall alignment." Note that HumRRO will not directly address overall alignment since Segment C was not administered operationally in year one (2018–19).

The content of the CAST will also rotate across years, each year sampling different content from the CA NGSS. The rotation is designed to allow CAST to address the full breadth of the CA NGSS in a three-year span. This alignment study was conducted during the first operational year of testing, so it will not be possible to evaluate how well CAST addresses the breadth of the content standards over three years. HumRRO will be able to use the initial year's data, however, to estimate whether one administration can address roughly one third of the intended PEs. If so, the three-year rotation is feasible as a sampling plan for addressing the full breadth of the CA NGSS.

These two ways of considering alignment match the CAST test blueprint design explicitly. The blueprint indicates, "For scoring and reporting purposes, each of the three science domains will constitute one third of the test (items written to assess PEs associated with Engineering, Technology, and Application of Science will be assigned to one of the three science domains, depending upon the context of their stimulus)." It continues, "For the segments contributing to individual student scores (Segment A and Segment B), it is not possible to assess all PEs in a single testing year. For example, there are 14 PEs assessed in grade five, each of which would require multiple items to fully assess. As a result, PEs assessed in Segment A and Segment B will be rotated from year to year so that all PEs can be assessed in the segments contributing to individual scores over the course of a three-year period." HumRRO will use student-level alignment results to evaluate the CAST for this purpose.

The blueprint describes the use of Segment C: "For the segment contributing only to group scores (Segment C), matrix sampling (the administration of a number of different versions across the state) will allow for assessment of all PEs annually at a state-wide level." An alignment study of the full assessment, including Segment C, will not be a part of this report.

Research Questions

Activities conducted for the CAST Alignment Study were designed to provide information to answer the following research questions:

1. To what extent do the test design and test blueprints for the CAST support the claims to be made about student performance on the assessment?

2. To what extent does the test blueprint for the CAST represent an appropriate sampling of the content as set forth in the CA NGSS?

3. To what extent do the CAST test forms and test items reflect the test design and test blueprints?

4. To what extent do CAST tasks and items integrate more than one disciplinary core idea, crosscutting concept, and/or science and engineering practice?

5. To what extent do CAST test forms show balance across the disciplinary areas used for scoring and reporting purposes (physical sciences, life sciences, earth and space sciences)?

6. Do the CAST items range from low to high cognitive complexity and provide a sufficient number of items across the range of cognitive complexity?

7. How well does CAST fit the population being tested, in terms of the distribution of item difficulties within test forms and the distribution of student ability?

Responses to these research questions provided in the forthcoming alignment study report should guide future item development and provide validity evidence for the CAST suitable for submission for federal peer review under ESSA. The next sections describe the methodology and progress to date.

Methods and Progress to Date

Component 1: Evaluation of CAST Contractor Documentation

HumRRO conducted an initial review of contractor documentation to evaluate how alignment was considered during test development. This review was guided by the *Standards for Educational and Psychological Testing* (APA, AERA, NCME, 2014), which describe requirements for developing, reviewing, and trying out test items as well as requirements for documenting the processes used. Prior to conducting the review of contractor documentation, HumRRO identified relevant standards (hereafter referred to as Testing Standards) and established a rating form to capture reviewers' ratings. The rating form was compared to the federal peer review guidance to ensure all relevant components were indicated on the form. Additional components were added to the form as necessary to ensure clear parallels to the federal peer review guidance. Testing Standards and rating components were selected to support the claims structure established for the CAST. HumRRO submitted a draft rating form to the CDE Contract

Monitor for review and approval in advance of the rating tasks. Table 3.1 presents the scale used for the review. Ratings were made for each identified Testing Standard.

HumRRO requested relevant documents from the testing contractor. These documents included item and form development guidance, test blueprints and item specifications, item tryout and review procedures, procedures for reviewing and addressing item tryout information, and validity and reliability evidence for the test forms. When validity or reliability evidence was not yet available, the testing contractor provided plans for data collection and analyses, as well as any criteria that will be used to judge the appropriateness of the assessments' alignment. HumRRO also requested any guiding documents that illustrate the overall goals and philosophy underlying the assessment, such as a theory of action, interpretive argument, or other similar documents. HumRRO downloaded these electronic documents from the testing contractor's secure FTP site.

For each identified standard or rating component, three HumRRO researchers, all experienced in third-party evaluation of assessment systems, independently assigned an overall rating for each of the relevant standards based on the evidence collected. For example, HumRRO researchers reviewed evidence related to Testing Standard 4.8: "The test review process should include empirical analyses and/or the use of expert judges to review items and scoring criteria. When expert judges are used, their qualifications, relevant experiences, and demographic characteristics should be documented, along with the instructions and training in the item review process that the judges receive."

After an initial round of independent ratings, the three researchers met to discuss areas of disagreement (non-adjacent ratings) and to identify any gaps in the documentation received. HumRRO then followed up with the testing contractor regarding questions and requested additional documentation. Once all standards were independently rated, the researchers compared and discussed their ratings to reach a final consensus rating for each standard or rating component. Table 3. Provides an example of the rating scale. Note that compound standards may be split into several component rating dimensions.

Table 3.1. Rating Scale for Evaluating Strength of Evidence for Testing Standards

Rating Level	Description
1	No evidence of the Testing Standard found in the documentation
2	Little evidence of the Testing Standard found in the documentation; less than half of the Standard covered in the documentation and/or evidence of key aspects of the Standard could not be found.
3	Some evidence of the Testing Standard found in the documentation; approximately half of the Standard covered in the documentation, including some key aspects of the Standard.
4	Evidence in the documentation mostly covers the Testing Standard; more than half of the Standard covered in the documentation, including key aspects of the Standard.
5	Evidence in the documentation fully covers all aspects of the Testing Standard.

As ratings were completed, researchers recorded their ratings as well as the document(s) containing the evidence from which each rating was made. A single rated Testing Standard may reference multiple documents. Raters provided a written justification for each rating, noting strengths and areas where evidence was potentially missing, undocumented, or incomplete. For any criterion receiving a “1” rating, HumRRO ensured that no evidence for the Testing Standard was found because no evidence exists, rather than having that rating result from a logistical or communication error in which the testing contractor did not provide the necessary materials. For consensus ratings of 1, HumRRO made an additional request to the testing contractor to verify evidence for that criterion was truly missing. If additional evidence was provided, HumRRO revised ratings as appropriate. The ratings were arranged by claim to facilitate validation. CAST has four claims, one overall and three separate science domain claims, which are included in table 3.2 (excerpted from the CAST blueprint). Bold type highlights domain-level claims.

Table 3.2. CAST Domain Claims

Domains	Description
3D Overall	Students can demonstrate performances associated with the expectations of the California Next Generation Science Standards, through the integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts **across the domains of Physical Sciences, Life Sciences, Earth and Space Sciences, and Engineering, Technology, and Application of Science.**
3D Physical Sciences	Students can demonstrate performances associated with the expectations **in the disciplinary area of Physical Sciences** within the California Next Generation Science Standards, through the integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts.
3D Life Sciences	Students can demonstrate performances associated with the expectations **in the disciplinary area of Life Sciences** within the California Next Generation Science Standards, through the integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts.
3D Earth and Space Sciences	Students can demonstrate performances associated with the expectations **in the disciplinary area of Earth and Space Sciences** within the California Next Generation Science Standards, through the integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts.

A draft report on Component 1 was provided to CDE on January 24, 2019. That report contained preliminary findings based on the initial review of CAST documentation provided by CDE and the testing contractor. Additional documentation was requested following that initial report and the draft is being updated and finalized for inclusion as a chapter in the CAST alignment report. The testing contractor provided additional documentation to enable us to make final ratings.

The next components for the study required input from panels of educators and cooperation from the test developer. HumRRO conducted a single workshop to gather panel input; Component 2 provides a description of the recruitment process and the makeup of the panels, as well as workshop plans for the panelists. We describe in Component 2 the documentation and access needed from the test developer to conduct the review.

Component 2: Panel Evaluation of CAST Item Content

CAST Alignment Workshop: Evaluating Items for Alignment to CA NGSS

HumRRO worked collaboratively with the CDE Contract Monitor and other staff to recruit and select a group of 18 educators to serve on three CAST alignment review panels (one elementary, one middle, and one high school panel). Panelists were very familiar with the CA NGSS and were required to have at least three years of experience as California educators. The CDE provided HumRRO with contact information for two sources for panelists: (a) the science subcommittee of the Curriculum and Instruction Steering Committee (CISC), comprised of 13 CA NGSS-knowledgeable educators for all grade levels, and (b) the CA NGSS Early Implementers, comprised of 26 educators for levels kindergarten through grade eight. In addition to content requirements (e.g. biology, chemistry, physics, earth science), each panel also included educators with experience teaching English learners (ELs) and/or students with disabilities (SWD) who take the CAST.

HumRRO secured the meeting space and arranged for meals during workshop days, arranged lodging and travel for panelists, and provided all necessary equipment for the workshop, including two laptop computers per panelist (one laptop to view the assessment items and the other to document their ratings). Panelists were required to sign nondisclosure agreements as a condition of participation.

HumRRO conducted a two-day workshop February 28–March 1, 2019, during which panels of educators evaluated how well each item assessed the CA NGSS. The CAST Alignment Study Director provided alignment training to the full group of panelists (see Appendix E for the content of the training presentation). Each HumRRO panel facilitator provided additional training to the grade level panel. CDE staff, including the contract monitor and Science Program staff, observed the workshop and were available to respond to HumRRO facilitators' and panelists' questions about the CAST.

During the alignment workshop, panelists made ratings regarding what science standards items assessed, accounting for the three-dimensional nature of the standards. They rated each item according to its cognitive complexity requirements. They discussed discrepant ratings and reached consensus or near consensus when they disagreed about important ratings. The data produced during the panel workshop will be used to evaluate alignment of CAST to the CA NGSS. Panelists also assigned a cognitive complexity (depth of knowledge or DOK) rating to each item. HumRRO used Webb's DOK rating system, to correspond to the cognitive complexity metric used by the testing contractor and provided in the item metadata. It was important the DOK

ratings used by panelists were consistent with those used by the testing contractor to allow for direct comparisons of the ratings.

For the workshop, the testing contractor provided HumRRO access to all test items in the same format they are viewed by students. This included items from Segments A and B. The testing contractor also provided HumRRO access to item metadata (i.e., item parameters, *p*-values, DOK, and coding to the CA NGSS for each science item). Panelists made their independent ratings, came to consensus, and then compared their ratings to the coding from the testing contractor. All ratings were initially made without seeing the metadata, but panelists were allowed to change their ratings if they judged the metadata to be a better fit than the initial consensus rating.

This approach to alignment was conceptualized to verify that CAST uses a reasoned approach to sampling the content within the science standards. HumRRO will not compute or report proportions of potential combinations of science standards addressed on the CAST (e.g., DCI x SEP x CCC). This would be inappropriate because the full potential breadth of the standards would not be possible to represent in a single summative assessment. It would also be inappropriate because the CDE has conceptualized the science standards based on representation of the PEs (rather than on all possible combinations of dimensions).

In order to capture the specificity of the CAST items, and to verify that CAST represents the intended blueprint, HumRRO used the CAST Item Specifications as the alignment guide. These specifications guided the creation of the CAST items and represent the intended CAST measurement construct. Panelists were given access to the specifications during the item rating processes.

The information panelists coded into the forms came from a coding guide provided to panelists based on CAST item development guidance. The codes represent the standards which each item was developed to measure. There are codes specific to SEP Sub-Practice Assessment Targets, DCI Assessment Targets, and CCC Assessment Targets. A sample of the codes for 5-PS1-1 Matter and Its Interactions is listed in table 3.3. The codes were arranged by domain (Life, Physical, and Earth and Space Sciences).

Table 3.3. Sample CA NGSS Standards Coding Information for Panelists

Dimensions	CA NGSS Standards
SEP Sub-Practice	2.1 Ability to develop models
SEP Sub-Practice Assessment Targets	2.1.1 Ability to determine the components as well as relationships among multiple components, to include or omit, a scientific event, system, or design solution
SEP Sub-Practice Assessment Targets	2.1.3 Ability to represent mechanisms, relationships, and connections to illustrate, explain, or predict a scientific event
DCI Assessment Target (Option for PS1.A.4)	PS1.A.4a. Develop a model of matter with microscopic particles as the components.
DCI Assessment Target (Option for PS1.A.4)	PS1.A.4b. Describe bulk matter as being composed of tiny particles of matter that cannot be seen.
DCI Assessment Target (Option for PS1.A.4)	PS1.A.4c. Describe the behavior of many tiny particles to explain observable phenomena involving bulk matter.
DCI Assessment Target (Option for PS1.A.4)	PS1.A.4d. Explain observable phenomena by using a model of bulk matter composed of many tiny particles.
CCC Assessment Target (Option for PS1.A.4)	Student can: CCC3 Identify that natural objects exist from the very small to the immensely large.

CAST Alignment Criteria

Alignment criteria were developed by HumRRO and reviewed by staff from the National Center for Improvement in Educational Assessment (Center for Assessment). The reviewers were highly experienced in both alignment methodologies and the CA NGSS. Reviewers made several comments that helped to clarify how the criteria would be communicated and operationalized for the study. The criteria were presented to California's CAASPP Technical Advisory Group (TAG) and finalized prior to the alignment workshop.

It is important to remember that no assessment is perfectly aligned. These criteria were developed based on the documentation provided by CDE and the testing contractor and they represent several aspects of the overall alignment of the CAST to the CA NGSS. Failure to meet any single criterion does not indicate that the test is invalid or flawed in some way, only that that aspect of the assessment may need to be addressed through future item development or by other means. An alignment study should be formative in nature and provide the state and the testing company with actionable results to make the assessment more closely mirror the CA NGSS.

The Webb alignment method (1997, 1999, 2005) was originally designed to align content standards with large-scale assessments. Dr. Norman Webb researched and refined this method over time. His approach is often cited and has been reviewed by the

Council of Chief State School Officers (CCSSO).[1] The Webb method includes four major indicators to evaluate alignment. These indicators rely on statistical analyses to assess how well items on the assessment, regardless of item type and point value, match the state's content standards. The four alignment indicators are: categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance-of-knowledge representation. While it was not appropriate to implement Webb's methodology for this study, mainly because of the multidimensional nature of the content standards and the way the content is sampled across years, we did use Webb's criteria to help guide our methodology and the development of criteria for judging the alignment of the CAST. Below, we briefly describe Webb's criteria, and then describe the similar criteria planned for use with CAST.

Webb's ***Categorical Concurrence*** is a basic measure of alignment between content standards and test items. This term refers to the proportion of overlap between the content stated in the standards document and that assessed by items on the test. Webb's criterion is based on the minimum number of items required to achieve acceptable reliability for reporting. HumRRO prefers to directly examine the reliability of the science assessments, which will be available in the forthcoming technical report[2] for the CAST. Reliability of scores should be evaluated for overall science scores and sub-scores at the student level and for any aggregate scores or sub-scores computed for schools, districts, or the state.

Webb's categorical concurrence criterion is derived by determining if there are at least six items per reporting category on the assessment. California will produce an overall student score and sub-scores at the domain level (e.g. life, physical, and earth and space sciences). So, at the most basic level, California could meet Webb's criterion if at least six items per domain were included on the assessments. This would not be a robust criterion for determining the sufficiency of items for generating reliable student scores.

The California science standards were written as performance expectations (PEs) through which students can demonstrate understanding of the content. These PEs were developed based on what students are expected to have learned at each grade level. The PEs incorporate DCI, SEP, and CCC. Test items might directly address the PE, or they might address the supporting DCI, SEP, or CCC. Ideally, an item would be linked to both a PE and some number of DCI, SEP, or CCC, but that may not always be possible given the relatively discrete nature of selected-response test items. It may be necessary to address all aspects of a standard through multiple test items.

For this criterion, HumRRO will report the proportion of items that panelists match to one of the targeted PEs for science. The proportions also indicate the number of items not judged to relate to any PE. To be judged acceptable, at least 50 percent of the test items should be directly matched to a PE. HumRRO will use a 50 percent match to PE as one component of this criterion because some items are expected to be matched only to DCI, SEP, or CCC. Ideally, all items would match at least one PE, DCI, SEP, or

[1] See https://ccsso.org/sites/default/files/2018-07/TILSA%20Evaluating%20Alignment%20in%20Large-Scale%20Standards-Based%20Assessment%20Systems.pdf for background information on alignment.

[2] The technical report will be authored by ETS.

CCC. However, it is possible for an assessment to have acceptable alignment with one or two weak items (as judged by panelists). To be judged acceptable for the second component of this criterion, at least 90 percent of items should be matched to either a PE, DCI, SEP, or CCC. To be judged acceptable, the test form must meet both components. HumRRO will refer to this criterion as **Link to Standards**.

In addition to the criteria, HumRRO will also report the full item-level data to CDE. This will include a side-by-side comparison of the panelists' final consensus data with the testing contractor's metadata. Items that are not linked to a PE or DCI will be flagged for scrutiny. Descriptive statistics will be used to indicate the number and proportion of items linked to PE, DCI, SEP, and CCC, as well as the number of items that are linked to the physical, life, and earth and space sciences.

Webb's ***Depth-of-Knowledge*** (DOK) ***Consistency*** statistic measures the type of cognitive processing required by items compared to the cognitive processing required by the matched content standards. For example, is a student expected to simply identify or recall basic facts, to use reason to manipulate information, or to strategize how to best solve a complex problem? For instance, a student may be asked to identify the planets of our solar system among several answer choices. This task should be less complex (have a lower DOK) than comparing the composition of the planets in preparation for landing unmanned probes.

The purpose of using DOK as a measure of alignment is to determine whether a test item and its corresponding standard are written at the same level of cognitive complexity. In Webb's method, panelists make two separate judgments about cognitive complexity, one rating for the standard and one rating for the item. These two judgments are compared to determine whether the item is written at the same level as the standard to which it is linked. Webb (1997) refers to this comparison as *Depth-of-Knowledge consistency*.

Webb's DOK consistency category is nearly impossible to implement when the standards are multi-dimensional. Doing so would require panelists to determine the DOK for each potential combination of standard and dimension. For science, it is also the case that the test standards can be interpreted in multiple ways and each combination of standard and dimension would represent a range of cognitive complexities depending on the specific knowledge, skills, and abilities that were being addressed. So, even if HumRRO could generate the number of DOK ratings required by the science standards, ratings would likely be vague, unreliable, and inflated (Webb's rule is to assign the higher DOK level if the standard is ambiguous). Therefore, no attempt will be made to match item DOK with standard DOK for this study.

It is still, however, important to determine if the CAST items reflect the level of cognitive complexity indicated by the CA NGSS. Looking at the standards more globally, HumRRO found they focused on requiring students to use their science knowledge and skills to investigate potentially unfamiliar phenomena. Focusing on science in this way means that students are expected to engage in more complex reasoning than simply recalling science terms or generating simple answers using familiar algorithms. HumRRO therefore reasoned that California's science assessment should include few, if any, low-complexity items. Webb uses a four-point scale for DOK. For an assessment based on

CA-NGSS, HumRRO would expect no more than 10 percent of items to be rated at level one. Webb's scale also includes a level four rating, which is seldom used for summative tests. This level of cognitive processing requires deep engagement of the students with the content, in multiple ways, typically over an extended period of time. This level is similar to producing a thesis or generating an extensive investigation of some scientific phenomenon a student would observe, collect data about, and generate a report to describe. HumRRO does not expect CAST to include level four items but does expect the assessments to be primarily a mix of DOK level two and three items. HumRRO also expects more level two items than level three items. Level three items require more input or time for students to respond, and it would not be practical to include primarily level three items on a summative assessment. In other states, notably Colorado, science standards are presented with a DOK range included. The range indicates the level of items and the level of instruction that are expected when addressing the content of the standards. In most cases in the Colorado standards, the range for standards is 1–3, with the mode clearly at level 2. HumRRO proposes setting California's DOK acceptability criterion such that no more than 10 percent of items are rated at level 1 and no less than 10 percent of items are rated at level 3. If there are more than 10 percent of items at level 1 or fewer than 10 percent of items at level 3, the DOK level of the items as a group would be judged too low to adequately represent the California science standards. HumRRO will refer to this criterion as **DOK Adequacy**.

In addition to the criteria, HumRRO will provide panelists' final DOK ratings for each item. These ratings will be compared to the DOK ratings included in the item metadata. Descriptive statistics (number and proportion of items at each level) will also be provided.

Webb's ***Range-of-Knowledge Correspondence*** examines the extent to which the test items reflect the full range of knowledge, skills, and abilities contained in the standards document. Where categorical concurrence notes whether a sufficient number of items on the test covers each general content topic (reporting category), the range-of-knowledge correspondence measure indicates the number of specific content objectives within each broader topic that are assessed by the test items.

Webb's range-of-knowledge correspondence criterion requires that at least 50 percent of the standards from each reporting category are addressed on the assessment. As stated above, California intends to report students' overall science scores and domain level sub-scores, but not finer-grained sub-scores (e.g. physics, chemistry, ecology, cross-cutting concepts, science and engineering practices). Meeting Webb's range-of-knowledge criterion would thus require that at least half of the full range of standards for science be represented on the tests. Given the three-dimensional nature of the standards, this criterion is not practical. The number of potential combinations of domains and dimensions represent too many standards to address in any single testing event. Even assessing at the PE level, if one were to address every PE on a single assessment, the number of required items would be impractical. HumRRO believes it is necessary, therefore, to sample the standards for assessing students. The standards emphasize students making meaning from information gathered from new or unfamiliar phenomena. They are expected to have a deep understanding of SEPs and CCCs, and that knowledge is expected to provide tools to use across DCIs in all content domains.

HumRRO will focus on SEPs and CCCs for this criterion rather than on trying to address the full breadth of the science standards.

Because students are expected to use their knowledge of SEPs and CCCs across multiple standards and content domains, HumRRO would expect these dimensions to be high priorities on California's science assessments. HumRRO also expects there to be few, if any, items on the tests that measure only an SEP or a CCC, and that these concepts are measured in context with DCIs from legitimate scientific phenomena. Items are coded to indicate if they measure an SEP or CCC, or both. HumRRO would expect at least 50 percent of the eight SEPs and seven CCCs to be directly measured by items on the tests. Hence, the assessments should contain items that address at least 4 SEPs and 4 CCCs to meet this criterion. HumRRO will refer to this criterion as **Range Adequacy**.

In addition to the criteria, HumRRO will provide descriptive statistics (number and proportion of items) matched to each SEP and CCC.

Webb's ***Balance-of-Knowledge Representation*** focuses on content coverage in yet more detail. In this case, the number of items matched to the content objective does matter. The balance of representation criterion determines whether the assessment measures the content objectives equitably within each content topic using only those content objectives identified by panelists as measured by the test item. Based on Webb's (1997) method, items should be distributed evenly across the objectives per content topic for good balance. The balance-of-knowledge representation is determined by calculating an index, or score, for each content topic. Each topic should meet or surpass a minimum index level to demonstrate adequate balance.

It would not be possible to compute a single interpretable balance-of-knowledge representation index for a three-dimensional assessment. The interaction of the dimensions and domains would yield too many objectives to include on a summative test form (and too many PEs). It does, however, make sense to consider that each content domain should be represented rather evenly, or purposefully, on an assessment. It would also be sensible to declare that the three dimensions should be represented rather evenly, or purposefully, on an assessment. Acceptability for California's science test will be determined using the same metric as Webb uses for balance-of-knowledge correspondence with the notable exception that it will be computed twice; once for domain, and again for dimension. Acceptability for each will be set at the same level Webb uses for traditional assessments (0.70). Both balance criteria must be met for the assessment to be considered adequately aligned. HumRRO will refer to this criterion as **Balance-of-Knowledge Correspondence (Revised for Science)**, or simply as **Balance**.

Finally, California's test items are written to be multi-dimensional. They are intended to measure more than isolated science content knowledge and are expected to address CCC and SEP in addition to DCI and/or specific PE. To address whether the items accomplish this goal, HumRRO will evaluate whether panelists agree that items are related to multiple science dimensions across DCI, CCC, and SEP. To be judged

acceptable, at least 90 percent of items should address more than one dimension. HumRRO will refer to this criterion as **Multidimensional Adequacy**.

Table 3.4 summarizes the criteria used to evaluate the alignment of CAST items to the California science standards. Failure to meet a single criterion would not indicate that the test is insufficiently aligned to generate meaningful scores, but that attention to that aspect of the test should be addressed through future item development. If several of the criteria were not met, it would signal that HumRRO should be concerned with the link between the assessment and the intended measurement construct.

Table 3.4. Criteria for the Alignment of Science Assessment Items to Standards

Criteria	Description
Link to Standards	Acceptable if 50% or more of the items are directly and clearly matched to a specific PE and at least 90% of items are matched to at least one PE, DCI, SEP, or CCC.
DOK Adequacy	Acceptable if fewer than 10% of items are rated as DOK level 1 and more than 10% of items are rated at DOK level 3 or 4 (using Webb's DOK definitions).
Range Adequacy	Acceptable if at least 50% of the CCCs and SEPs are aligned to test items (at least 4 CCCs and 4 SEPs)
Balance-of-Knowledge Correspondence (Revised for Science)	Webb's balance-of-knowledge correspondence criterion is used, computed for content domains and NGSS dimensions separately. Both must meet Webb's threshold of 0.70.
Multidimensional Adequacy	Acceptable if at least 90% of items are aligned to more than one dimension.

Next Steps

The CAST Alignment Study design, as presented in the 2018 Independent Evaluation Report, includes detailed descriptions of all activities HumRRO will complete for the study in its entirety. During 2019–20, HumRRO will complete analysis of data from the alignment workshop, update findings from the evaluation of contractor documentation relevant to select Testing Standards, and develop the CAST Alignment Study report for delivery to the CDE in spring 2020.

Chapter 4: California Alternate Assessment for Science Alignment Study Update

The alignment study for the California Alternate Assessment (CAA) for Science will aim to provide validity evidence for the CAA for Science as a measure of science achievement for the population of students for which it was designed—students with severe cognitive disabilities. This study will focus on links between the Science Core Content Connectors (alternate achievement standards, hereafter referred to as Science Connectors), and the test forms and test items developed to assess them. The Science Connectors are derived from the performance expectations of the California Next Generation Science Standards (CA NGSS), which also define the science construct(s) to be measured.

The CAA for Science is not a single end-of-year summative test but instead is designed to be administered following instruction throughout the school year. Four separate sessions, three operational and one field test, are administered each year, and each session consists of one embedded performance task (PT). Each operational PT addresses one science domain (i.e., life sciences, physical sciences, and earth and space sciences). Administration is not tied to an administration window, as for a typical summative assessment, and teachers will have discretion to administer each session when they have completed instruction on that specific domain during the school year. The students' performance on the three operational PTs will be aggregated to generate an overall science score at the conclusion of the school year. The CAA for Science is to be administered in grades five, eight, and once in high school. The high school assessment may be administered in grade ten, eleven, or twelve. Two Science Connectors are assessed in each PT, and each PT is expected to include two low and two medium complexity test items and one high complexity test item (numbers of score points will also vary by item). Each Science Connector has a corresponding set of five test questions prefaced by a nonscorable orienting activity designed to engage students with a science concept they were previously taught and will be assessed on.

As illustrated in figure 4.1, the Science Connectors are further broken down into more discrete focal knowledge, skills, and abilities (FKSAs) and essential understandings (EUs), which are basic concepts. Test questions are written to assess the FKSAs and EUs. There are one to six FKSAs and one EU for each Science Connector. Each EU but not all FKSAs for a Science Connector will be assessed in a single embedded PT.

Figure 4.1. CAA for Science standards continuum

Obviously, the full breadth of the Science Connectors available cannot be represented by three PTs that measure just two connectors each. There are 20, 24, and 28 Science Connectors for grades five, eight, and high school, respectively. The CAA for Science is expected to rotate connectors from year to year to build to fuller representation of the

content over time. All content from the 72 identified Science Connectors will be assessed in a five-year span.

Alignment studies for an assessment with this structure must approach evidence gathering in two ways. First, it must demonstrate that the aggregation of the three sessions provides an adequate representation of the science content specified by the Science Connectors. This alignment task supports the overall score and is the key evidence required by ESSA under federal peer review guidance. There is only one claim for the alternate assessment for science, and that claim indicates students should demonstrate performance "across the domains." Additionally, each session should adequately represent its tested domain, even if student-level scores are not produced at this level. Because teachers administer the assessment one-on-one, uneven or inadequate representation could lead to unwanted instructional or curricular changes over time. To avoid such consequences, test administrators should have confidence the assessment is a fair representation of the domain. While the sessions would not be expected to generate entirely reliable score estimates, each domain-level session should represent the intended domain. Data will be collected to demonstrate whether the Science Connectors are adequately represented, and those same data will be used to ensure the content domains are evenly represented.

It should be noted that any student-level results represent a sampling of the Science Connectors. The CDE will assess all of the Science Connectors identified in the CAA for Science blueprint over the course of five years. Adequate representation, as described above, means that each PT, one from each science domain (i.e., earth and space sciences, life sciences, and physical sciences), will assess the two Science Connectors per PT.

The research questions and methodology for this alignment study were designed specifically to address the structure and design of the CAA for Science and the results to be reported. The detailed design of the CAA for Science Alignment Study is included in the 2018–20 CAASPP Evaluation Plan, which was presented in the publicly available *2018 CAASPP Evaluation Report*. This chapter provides a brief summary of activities conducted to date in support of the study.

Research Questions

Activities conducted for the CAA for Science Alignment Study were designed to provide information to answer the following research questions:

1. To what extent do the test design and test blueprint for the CAA for Science support the claims to be made about student performance on the assessment?

2. To what extent do the test forms and test items for the CAA for Science reflect the test design and test blueprint?

3. To what extent do the CAA for Science PT items link to the Science Connectors?

4. How well do the CAA for Science PT items cover the range of cognitive complexity of the Science Connectors?

Responses to these research questions will be provided in a stand-alone alignment study report to be delivered to the CDE in 2020. The report should guide future item development and provide validity evidence for the CAA for Science suitable for submission for federal peer review under ESSA.

Methods and Progress to Date

Coordination with CAA for Science Test Contractor and the CDE

HumRRO's project manager and the CAA for Science Alignment Study director met with staff from the testing contractor (ETS) for CAA for Science and CDE staff via teleconference to coordinate activities for this study. Key topics discussed were (a) HumRRO's plans for data collection, (b) a review of CAA for Science assessment materials (e.g., online test content, Directions for Administration, planning guides), (c) documentation needed from ETS and CDE, (d) estimates of dates when files would be available to HumRRO from ETS, (e) panelist recruitment, and (f) process support needed for the alignment workshop (e.g., access to a secure online method for viewing items). As ETS was uncertain it could provide all materials and process support for the alignment workshop by August 30, 2019, HumRRO delayed the planned October 2019 workshop until November 2019.

For the workshop, ETS will provide access to all test items in the same format they are viewed by students and item metadata. For each science item, ETS will provide item parameters, *p*-values, cognitive complexity rating, and coding to the Science Connector, FKSA, and EU. HumRRO requested all materials and support from ETS be provided by October 1 to allow adequate time for workshop preparation.

After reviewing the publicly available CAA for Science blueprint, HumRRO created a list of questions for the CDE regarding the role Science Connectors, FKSAs, and EUs played in test development. The CDE's responses will inform HumRRO's design of the panelist data collection instrument for the alignment workshop.

Component 1: Evaluation of CAA for Science Contractor Documentation

HumRRO collected and began an initial review of contractor documentation to evaluate how alignment was considered during test development. This review was guided by the *Standards for Educational and Psychological Testing* (APA, AERA, NCME, 2014), which describe requirements for developing, reviewing, and trying out test items as well as requirements for documenting the processes used. Prior to conducting the review, we identified relevant standards (hereafter referred to as Testing Standards) and established a rating form to capture reviewers' ratings. The rating form was compared to the federal peer review guidance to ensure all relevant components were indicated on the form. Additional components were added to the form as necessary to ensure clear

parallels to the federal peer review guidance. Testing Standards and rating components were selected to support the claims structure established for the CAA for Science. Table 4.1 presents the scale that will be used for the review.

Table 4.1. Rating Scale for Evaluating Strength of Evidence for Testing Standards

Rating Level	Description
1	No evidence of the Testing Standard found in the documentation
2	Little evidence of the Testing Standard found in the documentation; less than half of the Standard covered in the documentation and/or evidence of key aspects of the Standard could not be found.
3	Some evidence of the Testing Standard found in the documentation; approximately half of the Standard covered in the documentation, including some key aspects of the Standard.
4	Evidence in the documentation mostly covers the Testing Standard; more than half of the Standard covered in the documentation, including key aspects of the Standard.
5	Evidence in the documentation fully covers all aspects of the Testing Standard.

Note. Compound Testing Standards may be divided into several component rating dimensions.

HumRRO requested relevant documents from ETS. These documents included item and form development guidance, test blueprints and item specifications, item tryout and review procedures, procedures for reviewing and addressing item tryout information, and validity and reliability evidence for the test forms. Because validity or reliability evidence was not yet available, HumRRO requested plans for data collection and analyses, as well as any criteria that will be used to judge the appropriateness of the assessments' alignment. HumRRO also requested any guiding documents that illustrate the overall goals and philosophy underlying the assessment, such as a theory of action, interpretive argument, or other similar documents. Table 4.2 lists the documents the testing contractor provided to HumRRO, or that were made available on the CDE website, to date.

For each identified Testing Standard or rating component, three HumRRO researchers, all experienced in third-party evaluation of assessment systems, will independently assign an overall rating based on the evidence collected. For example, HumRRO researchers will review evidence related to Standard 4.8, "The test review process should include empirical analyses and/or the use of expert judges to review items and scoring criteria. When expert judges are used, their qualifications, relevant experiences, and demographic characteristics should be documented, along with the instructions and training in the item review process that the judges receive." After an initial round of independent ratings, the three researchers will meet to discuss areas of disagreement (non-adjacent ratings) and to identify any gaps in the documentation received. We will

then follow up with the testing contractor to ask questions and request additional documentation as needed. Once all Testing Standards are independently rated, the researchers will compare and discuss their ratings to reach a final consensus rating for each Testing Standard or rating component.

Table 4.2. CAA for Science Documents Received to Date from Test Contractor

Document Focus	Document File Name
Describes how test forms are assembled.	CAA for Science Blueprint
Document produced by ETS psychometrics group listing the statistical parameters for individual items and the form as a whole.	2019–20 Statistical Specifications
Training materials for outside item writers consisting of a slide deck and handouts (multiple documents in a zip folder)	Item Writer Workshop materials (3 documents)
Training materials for teacher reviews of items, consisting of a slide deck and handouts	Item Review Meeting materials (5 documents, 1 slide deck, 1 spreadsheet)
Final configuration of the four test versions making up the 2019–20 administration, with the details of each of the four PTs that constitute a "version". (Science Connectors assessed, and item set status as operational or FT).	2019–20 Test Design documentation (2 documents)
Excel documents that contain assessment metadata for each PT, including but not limited to item number, sequence, Science Connector, item type, key, and statistical information. One per grade (Gr 5, 8, and high school).	2019–20 Form Planners (3 spreadsheets)
Online, self-guided training module that test examiners must complete in order to be certified to administer CAAs each year.	2019–20 Test examiner tutorial
High-level explanation of the 2019–20 CAA for Science administration for LEA coordinators and test examiners, including listing of the Connectors, by science domain, that are assessed this year.	2019–20 Admin Planning Guides
Item authoring template	2019–20 Directions for Administration

HumRRO has begun categorizing by Testing Standard the CAA for Science documentation provided by ETS. Table 4.3 lists relevant standards against which supporting documentation will be evaluated.

Table 4.3. Relevant Testing Standards for CAA for Science Documentation

Testing Standard
Standard 1.9. When a validation rests in part on the opinions or decisions of expert judges, observers, or raters, procedures for selecting such experts and for eliciting judgments or ratings should be fully described. The qualifications and experience of the judges should be presented. The description of procedures should include any training and instructions provided, should indicate whether participants reached their decisions independently, and should report the level of agreement reached. If participants interacted with one another or exchanged information, the procedures through which they may have influenced one another should be set forth.
Standard 1.11. When the rationale for test score interpretation for a given use rests in part on the appropriateness of test content, the procedures followed in specifying and generating test content should be described and justified with reference to the intended population to be tested and the construct the test is intended to measure or the domain it is intended to represent. If the definition of the content sampled incorporates criteria such as importance, frequency, or criticality, these criteria should also be clearly explained and justified.
Standard 1.12. If the rationale for score interpretation for a given use depends on premises about the psychological processes or cognitive operations of test takers, then theoretical or empirical evidence in support of those premises should be provided. When statements about the processes employed by observers or scorers are part of the argument for validity, similar information should be provided.
Standard 2.3. For each total score, subscore, or combination of scores that is to be interpreted, estimates of relevant indices of reliability/precision should be reported.
Standard 3.2. Test developers are responsible for developing tests that measure the intended construct and for minimizing the potential for tests' being affected by construct-irrelevant characteristics, such as linguistic, communicative, cognitive, cultural, physical, or other characteristics.
Standard 3.9. Test developers and/or test users are responsible for developing and providing test accommodations, when appropriate and feasible, to remove construct-irrelevant barriers that otherwise would interfere with examinees' ability to demonstrate their standing on the target constructs.
Standard 4.0. Tests and testing programs should be designed and developed in a way that supports the validity of interpretations of the test scores for their intended uses. Test developers and publishers should document steps taken during the design and development process to provide evidence of fairness, reliability, and validity for intended uses for individuals in the intended examinee population.
Standard 4.1. Test specifications should describe the purpose(s) of the test, the definition of the construct or domain measured, the intended examinee population, and interpretations for intended uses. The specifications should include a rationale supporting the interpretations and uses of test results for the intended purpose(s).

Table 4.3. (cont.)

Testing Standard
Standard 4.6. When appropriate to documenting the validity of test score interpretations for intended uses, relevant experts external to the testing program should review the test specifications to evaluate their appropriateness for intended uses of the test scores and fairness for intended test takers. The purpose of the review, the process by which the review is conducted, and the results of the review should be documented. The qualifications, relevant experiences, and demographic characteristics of expert judges should also be documented.
Standard 4.12. Test developers should document the extent to which the content domain of a test represents the domain defined in the test specifications.
Standard 12.4. When a test is used as an indicator of achievement in an instructional domain or with respect to specified content standards, evidence of the extent to which the test samples the range of knowledge and elicits the processes reflected in the target domain should be provided. Both the tested and the target domains should be described in sufficient detail for their relationship to be evaluated. The analyses should make explicit those aspects of the target domain that the test represents, as well as those aspects that the test fails to represent.

Component 2: Panel Evaluation of CAA for Science Item Content

Preparing for the CAA for Science Alignment Workshop

Beginning in April 2019, HumRRO worked collaboratively with the CDE contract monitor to recruit 18 educators to serve on three CAA for Science alignment review panels (one grade five, one grade eight, and one high school panel). Panelists are required to have a bachelor's degree and experience as a California teacher, to include experience working with severely cognitively disabled students or students with mild to moderate disabilities. Ideally, most members of each panel will also have familiarity with the CA NGSS and the Science Connectors. ETS provided HumRRO with contact information for potential panelists who had in some capacity supported development of the CAA for Science. HumRRO recruited from the ETS list, although educators who had been item writers for the CAA for Science were disqualified. HumRRO also provided text for the CDE to include in *Assessment Spotlight* (Issue 54), inviting interested teachers to contact HumRRO about this professional development opportunity.

After locating a venue in the Sacramento area, HumRRO set the date for the two-day workshop as November 4–5, 2019. HumRRO secured the meeting spaces, arranged lodging and travel for confirmed panelists, and began planning for all necessary materials, processes, and equipment for the workshop, including two laptop computers per panelist (one to view the assessment items and the other to document ratings).

CAA for Science Draft Alignment Criteria

CAA for Science alignment criteria will parallel the criteria developed for CAST. Draft alignment criteria for CAST were developed by HumRRO and reviewed by CDE's CAASPP Technical Advisory Group, the National Center for Improvement in Educational Assessment (Center for Assessment), and CDE staff. The reviewers were highly experienced in both alignment methodologies and the NGSS. Reviewers made several comments that helped to clarify how the criteria would be communicated and put into operation for the study. These criteria are the basis for the CAA for Science alignment workshop, but they will be modified to account for differences in the intention and structure of the CAA for Science and the expected inferences to be made from the scores.

It is important to remember that no assessment is perfectly aligned. These criteria are developed based on the documentation provided by CDE and the testing contractor and they represent several aspects of the overall alignment of the CA NGSS to the Science Connectors. Failure to meet any single criterion does not indicate that the test is invalid or flawed in some way, only that that aspect of the assessment may need to be addressed through future item development or by other means. HumRRO intends this alignment study to be formative in nature and provide the state and the testing company with actionable results to make the assessment more closely mirror the Science Connectors.

Next Steps

The CAA for Science Alignment Study design, as presented in the 2018 Independent Evaluation Report, includes detailed descriptions of all activities HumRRO will complete for the study in its entirety. During 2019–20, HumRRO will conduct the alignment workshop to collect item-level ratings from grade-level panels of educators, conduct analysis of data from the alignment workshop, collect additional contractor documentation and generate findings from the evaluation of all documentation relevant to select Testing Standards, and develop the CAA for Science Alignment Study report for delivery to the CDE in summer 2020.

References

Achieve. (2018). Criteria for Procuring and Evaluating High-Quality and Aligned Summative Assessments. Retrieved from https://www.achieve.org/files/Criteria03202018.pdf

California Department of Education. (2018a). Appendix A: Theory of Action for CAASPP and the Smarter Balanced Assessment System. In *California Assessment of Student Performance and Progress (CAASPP) 2018 independent evaluation report*.

California Department of Education. (2018b). *Integrating the CAASPP Tools to Create a Process of Improvement*. Presentation from 2018–19 CAASPP Institute workshops conducted across California.

Callingham, R., & Bond, T. (2006). Research in Mathematics Education and Rasch Measurement. Editorial in *Mathematics Education Research Journal*, 18, 2, 1-10.

Creswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Creswell, J.W., & Plano Clark, V.L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage Publications.

ETS. (2017). *California Science Test Blueprint*. Approved by the State Board of Education on November 8, 2017. (Revised 12/1/2017).

Darling-Hammond, L. & and Pechone, R. (2010) Developing an internationally comparable balanced assessment system that supports high-quality learning. *The National Conference on Next Generation Assessment Systems* (pp. 3–63). The Center for K–12 Assessment & Performance Management.

Flowers, C., Wakeman, S., Browder, D., & Karvonen, M. (2007). *Links for academic learning: An alignment protocol for alternate assessments based on alternate achievement standards.* Charlotte, NC: University of North Carolina at Charlotte. Retrieved from: http://www.naacpartners.org/LAL/documents/NAAC_AlignmentManualVer8_3.pdf

Hardoin, M. M., Thacker, A., Norman Dvorak, R., Becker, D. E. (2018). *California Assessment of Student Performance and Progress (CAASPP) Independent evaluation report* (2018 No. 087). Alexandria, VA: Human Resources Research Organization.

Marsh, J.A. (2012). Interventions promoting educators' use of data: Research insights and gaps. *Teachers College Record, 11*(2). 1–48.

Moss, P. A., & Haertel, E. H. (2016). Engaging methodological pluralism. In D. Gitomer and C. Bell (Eds.), Handbook of research on teaching (5th ed.), (pp. 127–247). Washington, DC: AERA.

National Research Council. (2014). *Developing Assessments for the Next Generation Science Standards*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18409.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and Practice* (4th ed.). Thousand Oaks, CA: Sage Publications.

Shepard, Hammerness, Darling-Hammond and Rust, 2007 [*citation to come*]

Smarter Balanced Assessment Consortium (2019), *Interim Assessments Interpretive Guide*, https://portal.smarterbalanced.org/library/en/reporting-system-interpretive-guide.pdf .

Weiss, R. (1994). *Learning from Strangers: The Art and Method of Qualitative Interview Studies*. New York, New York: The Free Press.

Wilson, M., & Draney, K. (2002). *A technique for setting standards and maintaining them over time*. In S. Nishisato, Y. Baba, H. Bozdogan, & K. Kanefugi (Eds.), Measurement and multivariate analysis (Proceedings of the International Conference on Measurement and Multivariate Analysis, Banff, Canada, May 12-14, 2000), pp 325-332. Tokyo: Springer-Verlag.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Glossary of Acronyms

Acronym	Gloss
AERA	American Educational Research Association
CAA	California Alternate Assessment
CAASPP	California Assessment of Student Performance and Progress
CA NGSS	NGSS for California Public Schools, Kindergarten through Grade Twelve
CAST	California Science Test
CAT	Computer-adaptive test
CCC	Crosscutting Concept (CA NGSS)
CDE	California Department of Education
DCI	Disciplinary Core Idea (CA NGSS)
DIF	Differential Item Functioning
DL	Digital Library
DOK	Depth of knowledge
EC	California Education Code
ECD	Evidence Centered Design
EL	English learner (student)
ELA	English language arts/literacy
ELPAC	English Language Proficiency Assessments for California
ESSA	Every Student Succeeds Act
ETS	Educational Testing Service
FIAB	Focused Interim Assessment Block
FKSAs	Focal Knowledge, Skills, and Abilities
FTP	File Transfer Protocol
GLT	Grade Level Team

IAB	Interim Assessment Block
ICA	Interim Comprehensive Assessment
LEA	Local educational agency
NCME	National Council on Measurement in Education
NGSS	Next Generation Science Standards
ORS	Online Reporting System
PE	Performance expectation (CA NGSS)
PII	Personally identifiable information
PLC	Professional Learning Community
PT	Performance Task
P-Value	Probability Value
SBE	State Board of Education
SEM	Standard Error of Measurement
SEP	Science and Engineering Practice
SE	Socioeconomically
SSR	Student Score Report
STEAM	Science, Technology, Engineering, Arts, and Mathematics
TAG	CAASPP Technical Advisory Group
TOSA	Teacher On Special Assignment

Appendix A: Impact Case Study 2018–19 Site Visit Protocol

Teacher Focus Groups

Who: Use this protocol for teacher focus groups at the elementary, middle, and high school levels. At elementary schools ALL teachers will answer questions for math and ELA. In middle and high school, there will most likely be a separate focus group for each content area.

Notes: Prior to each visit, please update the red text to include information appropriate to each focus group. Priority Questions are bolded. Depending on the pace after the first question, determine whether it makes sense to ask only bolded questions.

SCRIPT:

Good *[morning/afternoon]*. My name is *[HumRRO facilitator]* and this is [*HumRRO notetaker and POC for this LEA and its schools]*. We are with the Human Resources Research Organization, or HumRRO. Before we get started I'd like to make sure you are all aware that we will be recording today's focus group. This is for HumRRO internal purposes only, so that we can verify our notes are correct, and capture information we may miss. The recordings will not be shared with anyone from your school, LEA, or with the California Department of Education. Do you have any objections?

As a reminder, the Impact Study is not an evaluation of your school; it is an exploration of how CAASPP resources are used. HumRRO will keep information about participating LEAs, schools, and staff confidential.

HumRRO has collaborated successfully with many LEA and school leaders and school educators as part of our assessment evaluation work for the CDE for over 15 years, including the CAHSEE and the CAASPP. Your ability to inform us about how the current assessments and resources are functioning in the field is invaluable. Today, we are interested in hearing about your experience with various components of the CAASPP system—the summative assessments, interim assessments, and digital library resources. We are most interested in your use of these components during the 2018–19 school year; however, if you only have examples for questions from the 2017–18 school year, we would be interested in learning about those. We are particularly interested in learning about the context of your school, and how you use these components to impact classroom instruction and student learning. As you think about each question, please consider strengths and weaknesses about the program – we are interested in understanding what is working well, as well as where there is room for improvement.

In addition to information you provide today, we would also like to collect any materials or documentation you can provide to illustrate how you use CAASPP components. At the end of this interview, we will provide you with a list of examples of documentation we would be interested in reviewing. You will be asked to send these to your HumRRO POC *[name]*.

Before we jump into questions, can you please each share a bit about yourself – including your name, the grade and content you currently teach, and how long you have been teaching at this school, and total teaching experience overall. Your names will not be shared with others outside this group.

Name	Grade(s) and Subject(s)	Experience total	Experience at this school
N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A

We hope to get through several questions today and have limited time. As we go along, we understand that some answers to questions may be consistent across educators; therefore, if you have a similar response to a colleague a simple "agree" will do. We will try to start each question with a different teacher so that we can hear from each of you.

First, I want you to think about how you used the 2018 Smarter Balanced *[ELA/math/ELA and math]* summative assessment results.

1. **Can you describe how, when, and what types of summative assessment results are shared with teachers in your school?** *[Research Question 7]*
    a. **Can you explain any guidance or instruction you are given from LEA or school leadership on using summative assessment results, or describe any process in place for interpreting and reflecting on results?**
    b. **Can you describe a situation in which school leadership or you as a classroom teacher have modified materials, programs, or instructional practices, based in part on your understanding of summative results?**
    c. *[Elementary school only]* Can you describe any differences in the process for interpreting and using ELA summative assessment results versus the process for math results?

Next, let's discuss your use of interim assessments in the classroom. We are primarily interested in your use of Interim Assessment Blocks (IABs), but if you also use results of Interim Comprehensive Assessments (ICAs) to inform instruction, think of them, also, as you address questions. Please specify whether your responses refer to IABs, ICAs, or both.

2. **Are you able to make decisions about use of interim assessments – such as which ones are given and when – or are these decisions made at the LEA or school level?** *[Research Question 9]*
    a. **Can you describe any guidance or instruction you have been given on how and when to use interim assessments from school or LEA leadership?** *[Research Question 2]*
    b. Do you ever modify how IAs are administered to better meet your students' needs (e.g., for special needs students, English learners), classroom instruction, etc.? Explain.
    c. **Can you explain how well the particular IAs you are giving align to your instruction/curricula?**
    d. Do you ever administer IAs only to some students, but not others?
        i. If yes, can you provide an example of a time that lead you to decide to administer an interim assessment to an individual student or selection of students?
    e. Do you ever give the same IAB more than one time to the same student(s)? If so, explain your reasoning and how you make this decision.

3. **Can you give an example to describe how IA results are used in your classroom and at your school to help evaluate the effectiveness of practices for teaching the targeted standards?** *[Research Question 11]*
    a. Can you describe guidance that has been provided from school or LEA leadership concerning how to use information in this way?
    b. Do you think that student performance on IAs is a good indicator of the effectiveness of teachers' teaching practices?
    c. Who has a role in helping to evaluate whether teaching practices are effective? (your peers, principal, content specialist?)
    d. How might recommendations to adjust teaching practices be made, using information from IABs?

4. Can you provide an example of how you have used information from interim assessments to help track individual student progress toward grade-level goals? *[Research Question 12]*
    a. Would you say the current interim assessment results provide sufficient information for tracking individual student progress, or how might they be improved?
    b. Do you share information with your students to help them monitor their performance?

5. Can you provide an example of how you have used information from interim assessments to inform classroom instruction? *[Research Question 12]*
    a. Would you say the current interim assessment results provide sufficient information for informing classroom instruction, or how might they be improved?

Next, let's talk about your use of the Digital Library. This includes the instructional resources, professional learning resources, and playlist resources.

6. **Can you describe the types of Digital Library resources you use and why?** *[Research Question 13]*
    a. Can you describe any guidance or instruction from school or LEA leadership on what DL resources to use or how to use them?
    b. How often do you use them?
    c. How do you use information from the DL to assist with instruction, curricula, and assessment choices?

For our final questions, please consider your use of the full CAASPP system together – the summative assessments, interim assessments, and Digital Library.

7. Can you describe how you integrate results from the summative assessments, interim assessments, and formative assessments from the Digital Library (if used) with each other and with other measures to enhance classroom instruction and student learning? *[Research Question 5]*
    a. Can you explain any guidance provided at the LEA level for using results?
    b. Is this a fully integrated system or several related parts? If substantially integrated, how did this system evolve and who was involved?
    c. *[Elementary school only]* Is this coordinated across classrooms / schools?
    d. Do you face any challenges in using results for these purposes? Explain.
    e. Is this consistent for ELA and math?

8. **Can you describe how classroom assessment has changed since the onset of CAASPP, particularly the interim assessments**? *[Research Question 10]*
    a. Have CAASPP interim assessments replaced traditional classroom assessments? To what extent?
        i. How does the effectiveness of the CAASPP IABs compare to other classroom assessments you have used?
    b. Can you describe the impact of these changes on instruction and student learning?
        i. Have you been satisfied with these changes? Why or why not?

9. In your opinion, what aspects of CAASPP implementation have been most beneficial for improving classroom instruction and student learning in your classroom? [Research Question 3]

    a. Can you describe any aspects of CAASPP that could be improved to better inform classroom instruction and/or promote student learning in your classroom?

10. Is there anything else you would like to share about CAASPP components—strengths or weaknesses— that we have not covered in our questions?

Thank you so much for your participation today. Your input is highly important for better understanding the CAASPP system. As indicated, I will now pass around a list of examples of documentation we would be interested in obtaining to help us understand CAASPP use at your school. This document also provides an e-mail address where you can contact HumRRO staff involved in this project if you have any questions about our study. Going forward, over the course of the school year, your school POC [name] may be reaching out to you for additional information to provide HumRRO during monthly polling activities.

This page is intentionally blank.

Appendix B: Impact Case Study Polling Questions

School-Level

December 2018:

1. Can you describe examples of discussions about CAASPP you have had with administrators or teachers after their participation in interviews or focus groups?

January 2019:

2. Describe how educators at your school have used information from an IAB administered during the first semester of the 2018-19 school year to help make an instructional decision. Describe the goal of giving the IAB; include the grade level and a description of class characteristics. Evaluate how effective the IAB was as a tool for meeting the goal. Include as much detail as possible, including:

    a. How did the teacher prepare students to take the IAB?

    b. What administration settings were involved (designated supports or accommodations settings tried out)?

    c. Was it scored, or administered in a nonstandard manner?

    d. Were scores accessed in the IA Reporting System, and if so, which scores were reviewed?

    e. Were scores shared with students?

2b. (ALTERNATE VERSION FOR THOSE WHO DID NOT USE IABs) Conversely, if educators did not administer an IAB during the first semester or did not use information obtained from an IAB to help make an instructional decision, please explain why.

3. Describe educators' use or non-use of the Digital Library so far this school year.

If teachers <u>have</u> used at least one Digital Library resource:

3a. Describe when, which resources (names or URLs), and how they were used. For example, were they used as a formative assessment? to prepare students for an IAB or classroom assessment? to fill in gaps in the curriculum?

If teachers have <u>not</u> used at least one Digital Library resource:

3b. Describe what, if any, online CCSS-aligned ELA or math resources, targeting specific ELA or math standards, were used to help with lesson planning and formative assessment. Explain why the above resources were used instead of those provided by the Digital Library.

February 2019:

4. Please identify one educator for each IAB listed below and ask that they describe their use of the following IABs by responding to the sub-questions a–i below**.**

School	**Grade**	**Subject**	**IAB Name**
SCHOOL **NAME**	*IAB GRADE*	*IAB SUBJECT*	*IAB NAME (selected specific IABs unique to each school based on counts from data)*

a. Did you as an educator select this IAB, or was it a decision made at the school- or district-level? Explain.

b. If you were the one to select the IAB can you explain what made you select this particular one? If you did not select it, do you understand the justification for your school or district making the choice?

c. What were your goals for administering the IAB?

d. How did you administer the IAB (standardized or other manner such as whole class viewing items together, students working in pairs)?

e. How did you review and evaluate IAB results? Did you access results in the IA Reporting System, and if so, what information did you review?

f. Did you access the Digital Library Connections for the IAB to find instructional resources tied to the IA results? Explain.

g. How have you used the results from the IAB?

h. Were all of your goals achieved? If not, what could have been improved?

i. (If there was a hand scoring component) Please describe your process for scoring the hand-scored portion of the IAB. Were there any barriers that impacted your ability to score effectively? Did you feel you had the proper training and information to score accurately?

5. How is your school using CAASPP resources to prepare students for the Smarter Balanced ELA and mathematics summative assessments?

a. Is your school using or planning to use interim assessments (IABs or ICAs) in preparation for the summative assessment? If so, what are the goals for their use (e.g., learning which skills need targeted reteaching, increased exposure to CAASPP item types, practice using technology, trying out settings for designated supports, etc.)? If your school is using both IABs and ICAs, what are the goals for using each?

b. Is your school using or planning to use the Smarter Balanced ELA or math practice test or training test? If so, for what grade(s) or grade bands and content area(s)?

c. Is your school using or planning to use Digital Library resources to prepare students for the summative assessment? If so, what specific resource(s) are being used, and why did you decide to use those specific resources?

d. What do you consider the most valuable CAASPP resource to prepare students for the Smarter Balanced ELA and mathematics summative assessments? Why do you think this is the most valuable CAASPP resource for your students?

March 2019:

- *HumRRO provided documents and instructions for schools to conduct student focus groups. No new polling questions were asked.*

April 2019:

6. To enhance your classroom instruction throughout the year, how do you use and integrate the different measures of student achievement available to you (e.g., CAASPP Smarter Balanced summative and interim assessments, LEA-level or other benchmark assessments, classroom assessments, formative assessment resources in Digital Library, classroom activities)?

a. Describe what tools you use to diagnose student needs and measure student progress.

b. Explain with a few concrete examples how the different tools complement each other.

c. Describe any challenges you face when putting the parts together to get a picture of a student's/class's skills and abilities or their growth over time relative to the standards.

d. Do you feel the measures you use are adequate for understanding student achievement for purposes of guiding instruction? If not – what information is missing?

e. Are you required to give any assessments that you believe have limited value? For example, measures that reduce instructional time but provide little usable information.

May 2019:

7. Please describe how school-wide plans regarding the use of CAASPP components identified in a. – c. below are expected to be revised or continued during the 2019-20 school year. For each of the following, explain the rationale

for revising plans or "staying the course," based on what your LEA and/or school learned about how CAASPP components impact instruction and student learning during this current school year.

   a. Digital Library

   b. IABs or ICAs, including hand scoring

   c. caaspp.org online resources, including Practice and Training tests and teacher training videos/webcasts/workshops

8. Please poll your teachers (for middle school and high school, just ELA and math teachers) about any changes in the use of the use of CAASPP components identified in a. – c. below they are planning for their classes or grade-level, apart from those mandated by the school or LEA. For each of the following, their responses should explain the rationale for revising their plans or "staying the course," based on what they learned in their classrooms this year about how CAASPP components impacted instruction and student learning.

   a. Digital Library

   b. IABs or ICAs, including hand scoring

   c. caaspp.org online resources, including Practice and Training tests and teacher training videos/webcasts/workshops

LEA-Level

December 2018:

1. Can you describe examples of discussions about CAASPP you have had with administrators or teachers after their participation in interviews or focus groups?

January 2019:

2. Did you or someone else at your LEA provide anyone access to the IA Reporting System for IAB scores so far this school year? If so, who (position, school)? If not, did the scores get put into LEA's student information system, and what communication about IAB scores did LEA staff provide to schools, including teachers?

3. Including yourself, explain how staff at the LEA level have used the Digital Library resources so far this school year—either for themselves (e.g., to familiarize themselves with what school staff has access to) or to provide guidance to others. If LEA level staff have not used the Digital Library resources, explain why.

February 2019:

4. How is your LEA using CAASPP resources to prepare schools for the Smarter Balanced ELA and mathematics summative assessments?

    a. Describe any steps your LEA has taken to ensure schools at your LEA have the capacity to administer the summative assessments.

    b. Describe steps taken to ensure schools have the technology capacity.

    c. Have you provided any guidance or instruction on ensuring appropriate designated supports and/or accommodations are in place?

    d. Have you, or other staff at your LEA, mandated or encouraged the use of IABs and/or Smarter Balanced practice tests to prepare for the summative assessment? If so, what are your intended goals of their use? (e.g., learning which skills need targeted reteaching, increased exposure to CAASPP item types, practice using technology, trying out settings for designated supports, etc.)

    e. Have you, or someone at the LEA, suggested use of Digital Library resources to prepare students for the summative assessment? If so, what has been your guidance for doing so?

    f. What do you consider the most valuable CAASPP resource to prepare students for the Smarter Balanced ELA and mathematics summative assessments? Why do you think this is the most valuable CAASPP resource for your schools?

April 2019:

5. What guidance does your LEA provide to your schools for using and appropriately integrating different measures of student achievement in ELA and mathematics (e.g., CAASPP Smarter Balanced summative and interim assessments, LEA-level or other benchmark assessments, classroom assessments, formative assessment resources in Digital Library)?

    a. Name and describe the tools and resources you make available to schools for measuring student achievement for each of these purposes:

        - Accountability
        - Diagnose student needs
        - Inform instructional decision making
        - Program evaluation (if applicable)

    b. Describe (or provide documentation for) how the district recommends schools use the various pieces of assessment information, either in combination with each other or separately, for their intended purposes.

c. What challenges do you see teachers and/or school leaders facing when drawing conclusions about student achievement? Why do you think these challenges exist?

d. Has your district conducted a review of its local student assessment system to evaluate how it supports improved instruction and student learning? What (if any) changes or improvements would you like to see in your local assessment system?

May 2019:

6. Please describe how plans at the LEA level regarding the use of CAASPP components identified in a. – d. below are expected to be revised or continued during the 2019-20 school year. For each of the following, explain the rationale for revising plans or "staying the course," based on what your LEA learned about how CAASPP components impact instruction and student learning during this current school year.

    a. Digital Library

    b. IABs or ICAs, including hand scoring

    c. Other CAASPP resources such as caaspp.org online Practice and Training tests

    d. School support for implementing the full range of CAASPP components, such as professional development sessions, guidance documents, and trainings (including online CAASPP videos/webcasts)

Appendix C: Detailed LEA-Specific Findings for the Impact Case Study

Table of Contents

List of Tables

Appendix C: Detailed LEA-Specific Findings for the Impact Case Study

The following sections highlight findings for each study LEA. Each section begins with a table outlining the data used to generate the results. This is followed by a discussion of the characteristics of the LEA and its participating schools. Next are descriptions of the LEA's and the schools' use of Smarter Balanced Summative Assessments, Interim Assessments (IAs), and the Digital Library (DL). The description about each LEA depends on the information provided to HumRRO throughout the study period. As noted in Chapter 2, the LEAs and schools in the Case Study varied in their degree of participation in data collection activities, and some unique attributes or uses of CAASPP components resulted in inconsistent information.

To keep LEAs and schools anonymous, LEA and school codes are used for identification purposes. Each LEA is numbered (LEA-1 through LEA-7). Each school within an LEA includes that LEA code and an additional code based on the school level. Elementary schools are denoted "ES," middle schools "MS," and high schools "HS." Among the LEAs studied, there were some variations in the grades at each school level. For example, some elementary schools had kindergarten through grade five, while others also had a grade six. For the study, HumRRO classified findings from schools consisting of middle grades between elementary (ES) and high school (HS) as middle schools (MS), though some were named junior high schools. All the schools classified as middle schools included grades seven and eight, with some variation in the lowest grade.

We caution that these findings are based on the perspective of a small number of teachers within a small number of schools in a small number of LEAs. In addition, some of the responses represent limited understandings or awareness about the capabilities of the CAASPP System. We also note that some of the concerns expressed by teachers are those the CDE has already begun to address (see the Planned CAASPP System Updates in the Executive Summary and chapter 2).

LEA-1 Findings

LEA-1 participation included two elementary schools and one middle school. Table C1 summarizes HumRRO's analysis of the qualitative data gathered for this LEA. As shown, LEA-1 contributed greatly to this study, collaborating with its study schools throughout the first evaluation year.

Table C1. Summary of Data Sources for LEA-1

Data Source	Participants/Description
Site Visit Educator Focus Group	ES1 – 9 teachers in two focus groups: grade four (4), five (2), and six (3) ES2 – 6 teachers in one focus group: grade three (1), four (1), five (1), six (1), and multi (2) MS – 10 teachers in two focus groups: grade seven ELA (1), eight ELA (3), multi ELA (1); grade seven mathematics (2), and eight mathematics (3)

Table C1. (cont.)

Data Source	Participants/Description
Site Visit Leader Interview	ES1 – School POC (Principal) ES2 – School POC (Principal) MS – School POC (Principal) LEA – POC (CAASPP Coordinator), Superintendent, Director, Alternate CAASPP Coordinator, TOSA*(2), one interview
Monthly Polling	ES1 – December (POC), January (16 teachers); February (POC, 2 teachers); April (4 teachers); May (POC, 5 teachers) ES2 – December (POC); February (5 teachers); April (3 teachers); MS – January (12 teachers); February (POC, 3 teachers); May (4 teachers) LEA – December (POC, TOSA*), January (POC), February (POC)
End-of-Year Virtual Focus Groups	ES1 – POC ES2 – POC MS – POC LEA – POC
Student Focus Groups	ES2 – 9 students MS – 7 students
Documentation	• Interim Assessment information from CDE provided to educators

*Teacher on special assignment.

LEA and School Characteristics

LEA-1 is a medium-sized district in central California. The district includes 16 elementary schools and four middle schools. Table C2 summarizes the demographic characteristics and academic achievement of the LEA and its three participating case study schools. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest).

In 2017−18, the vast majority of students were socio-economically disadvantaged and there was a high percentage of English learners, particularly at the two elementary schools where more than half the students held this classification. The study schools were similar to the LEA in terms of academic achievement, with the majority of students not meeting or exceeding the grade level standards for ELA and mathematics. In 2017−18, LEA-1-ES1 performed best academically, with more than 40 percent of students having met or exceeded standards for both subject areas.

The LEA-1 schools were very similar in terms of student demographic characteristics. Most students at LEA-1-ES1 were below poverty level and many students at LEA-1-ES2 came to school with "a lot of baggage from home", as reported by the school POC. LEA-1-MS had many Spanish speaking students, with a growing number of Mixtec speakers from Oaxaca, Mexico. This school also had a large number of long-term English learners. The LEA-1 participating middle school reported that all students at the school received free lunch and the majority lived in the community and walked to school.

Table C2. Demographic Characteristics of LEA-1 and Its Participating Schools

Variables	LEA-1	LEA-1-ES1	LEA-1-ES2	LEA-1-MS
Enrollment	17,122	987	313	830
% Socioeconomically Disadvantaged	89%	95%	93%	97%
% Students with Disabilities	8%	7%	12%	10%
% English Learners	58%	57%	61%	44%
% Reclassified Fluent English Proficient	21%	23%	13%	47%
% Met or Exceeded ELA State Standards	31%	46%	25%	30%
% Met or Exceeded Math State Standards	23%	42%	23%	23%

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA overall and for each of the LEA's schools in the study. The second column (from top to bottom) shows in LEA-1 there was a total enrollment of 17,122 in 2017–18. Of these students, 89% were socioeconomically disadvantaged, 8% were students with disabilities, 58% were English learners, and 21% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 31% of students met or exceeded ELA state standards and 23% of students met or exceeded math state standards.*

The district provides an English Language Director at each school. In addition, staff from each school mentioned the availability of intervention teachers and other staff to provide remediation and support to students in addition to general classroom teachers. LEA-1-ES1 mentioned multiple interventionists and a full-time science, technology, engineering, arts, and mathematics (STEAM) teacher. LEA-1-ES2 noted they have bilingual aids, teachers who do one-on-one tutoring, a school counselor, and two outreach consultants who help kids with social skills. All office staff at LEA-1-MS speak Spanish, and they have a bilingual community liaison available to assist with parent-teacher communication. LEA-1-MS staff also includes a three-tiered Positive Behavioral Intervention and Support (PBIS) team to accommodate various levels of student needs.

The teachers participating in the focus groups across LEA-1 schools generally had many years' experience teaching at their current school, particularly at LEA-1-ES1 where more than half of the nine teachers had more than 20 years at the school. One teacher noted that longevity was not limited to the focus group participants; many teachers across the school had been at the school more than 20 years. Many of the LEA-1-MS teachers indicated additional roles beyond classroom teaching. For example, those participating in focus groups provided one-on-one tutoring and/or led extracurricular activities/clubs.

Various challenges were discussed at LEA-1-MS. Particularly, there was a change in 2017–18 that moved the school schedule from four 80-minute periods to six 50-minute periods. In addition, EL students lost the opportunity to take electives—classes like band, art, or technology—because they were put in an additional English Language Development (ELD) period. Some staff were trying to figure out how to schedule electives back in for these students. An increasing student population was discussed as another challenge at LEA-1-MS. One teacher mentioned classrooms with 36 students, despite an intended cap at 35. The teacher break room was recently converted to a classroom because more space was needed – so teachers take their lunches outside.

Professional Learning Communities and Professional Development

PLCs had a strong presence at the three LEA-1 schools, with time provided during school hours for the teacher-groups to meet. Teachers and school leaders across the three schools indicated the LEA-1 PLCs often were grade and/or subject specific and included discussions about instructional priorities, including essential standards; instructional strategies; and formative and interim assessments. PLC sessions often consisted of only the member teachers but occasionally included outside presenters or school leaders to support them with specific goals.

LEA-1-ES1 teachers indicated a strong team-oriented environment. Teachers worked together in grade-level teams as well as vertical teams designed to communicate across grade-levels. One teacher indicated PLCs had been built into LEA-1-ES1's culture from its start, with highly engaged and dedicated teachers. The school POC indicated teachers voluntarily familiarized themselves with the CCSS when California became a member of the Smarter Balanced Assessment Consortium. Teachers at LEA-1-ES1 were encouraged by school leaders to break down data to examine specific skills, and to use this information to better understand what students know and can do. Teachers noted a very open-door culture at the school. During a focus group, teachers expressed that they felt free to observe another teacher teaching a topic at any time to help them with their own teaching.

LEA-1-MS teachers expressed a very positive environment for collegiality between teachers. One noted that the teachers had a strong bond and they supported each other. Another noted they had conversations about their students and how they were doing behaviorally and academically in each other's classrooms. The teachers believed the closeness among teachers was something the students noticed and that it had a positive impact on them.

Teachers at LEA-1-ES1 participated in professional development opportunities, CAASPP-related and otherwise. Three teachers attended the CAASPP Summer Institute in 2018, and most staff attended the PLC Institute led by Solution Tree once in the past five years. The school POC noted an intent to have more teachers attend the CAASPP Summer Institute in the future. The LEA-1-MS school POC indicated some staff had participated in professional development related to CAASPP, though they were unclear about details. They noted information learned through training would typically be shared with others during PLC sessions.

Initiatives and Resources

Distributive practice is a process of generating daily, short instruction sessions. In LEA-1 distributive practice instructional sessions were developed for mathematics and ELA based on information learned from IABs and/or formative assessments built in SchoolCity[3]. The distributive practices are often developed to match the essential standards of the CCSS currently of focus. Essential standards are described as standards identified (by educators, leaders) as those that teachers need to focus on most heavily in instruction. HumRRO learned from one teacher at LEA-1-ES1 that the district embraced ideas presented by the professional development company Solution Tree, particularly the belief that establishing essential standards is the first step to providing a viable curriculum. School POCs and teachers across LEA-1 frequently created and used distributive practice as an instruction activity.

Finally, Achieve 3000, IXL, Aspire and StoryWorks Magazine were identified as additional resources used to enhance student learning related to the CCSS. LEA-1 adopted StudySync as the ELA curriculum, which included many digital resources aligned to the ELA standards and literature in each unit. At LEA-1-ES2, teachers mentioned using GoMath online curriculum for review purposes. LEA-1-MS teachers mention using MobyMax and Desmos for math.

LEA staff participating in interviews indicated LEA-1 attained funds for a big literacy initiative, providing professional development in the form of literacy coaching, for those teachers on special assignment (TOSAs) during the 2018–19 school year. Though each TOSA had a specified role for a grade level and/or content area, their roles expanded far beyond that, as they sought to build capacity in teachers across the district.

LEA-1-ES1 was successful in encouraging parental involvement, resulting in strong attendance at site council meetings. The meetings included a follow-up “coffee club,” with typically 10−15 parents in attendance.

Technology

LEA-1 schools indicated good access to technology. In fact, none of the three schools felt technology was a barrier to students doing well on the summative assessments. Similarly, teachers did not report issues with access to technology when administering online assessments. LEA-1-ES1 noted they are a one-to-one technology school; in 2018–19, they pilot tested a program in grades 1 and 3 for take-home Chromebooks.

[3] SchoolCity was purchased by Illuminate in June 2018. In this report we use the name used by the LEA.

Use of CAASPP Components

According to LEA staff participating in the LEA interview, the 20 schools in this LEA are on very different paths regarding:

- use of Interim Assessment Blocks (IAB) and Digital Library (DL),
- use of data from assessments to drive instruction,
- use of PLCs,
- understanding the CCSS, and
- expectations of student knowledge and ability.

The three study schools were further along in incorporating CAASPP resources in 2018–19 compared to other schools in the LEA. According to LEA staff, the two participating elementary schools incorporated CAASPP components in different ways. Specifically, LEA-1-ES1 used CAASPP components for intervention and filling in gaps, whereas LEA-1-ES2 used the CAASPP components in regular instruction.

The district used SchoolCity in addition to CAASPP components to generate and administer formative, interim, and diagnostic assessments. Teachers across the LEA noted that SchoolCity assessments were easy to create, and generated data that indicate which students need to be retaught a concept and which have mastered a skill. Teachers and school leaders indicate SchoolCity has a large item bank and is user friendly. School leaders and teachers frequently brought up use of SchoolCity. The degree to which a school, or grade within a school, chooses to use SchoolCity assessments or CAASPP interim assessments varied, with some electing to use both frequently and others choosing one more heavily over the other. The LEA-1-MS school POC liked that it was possible to modify items in SchoolCity, something the CAASPP IABs do not allow. Though many teachers and school leaders indicated satisfaction with SchoolCity, the staff of one elementary school preferred not to use SchoolCity and wished they could use CAASPP IABs only.

A LEA TOSA who had once worked at LEA-1-ES1 indicated that all staff at the school have always looked at assessment data to improve student learning and achievement, even before CAASPP. The LEA-1-ES1 school POC echoed this, stating that the teachers were very forward thinking and had always used data to understand what students know and can do.

Summative Assessments

LEA-level support was provided by the LEA CAASPP Coordinator prior to the testing window to ensure Chromebooks were current and had the required secure browser. The IT coordinator ensured district-wide bandwidth to support the testing window. Early in the school year, the LEA prepared a timeline to ensure a successful testing window. This included training new school test support coordinators, followed by training all test coordinators to ensure all students received proper supports. Each school had one school test coordinators who received district level training and were required to meet with all test administrators at their schools to train them on testing procedures. School leaders confirmed the LEA-level support and indicated through polling responses that they were

prepared to offer the accommodations and supports, and to administer the assessment. In addition, the IT staff at LEA-1-MS tested the school's bandwidth prior to testing.

The school POCs indicated various steps were taken to prepare students for the ELA and mathematics summative assessments. For example, both elementary schools indicated they practiced using IABs. LEA-1-ES1 mentioned using IABs to familiarize students with all the tools through IABs, practice, and training tests. LEA-1-ES2 mentioned use of IABs to expose students to performance tasks. LEA-1-MS middle school reported that the process of taking IABs led to student comfort with the summative exam.

Access to and Use of Summative Data

LEA-1 and its schools accessed summative assessment data in different ways. At the LEA-level, the CAASPP Coordinator pulled summative data from the Online Reporting System (ORS) and imported these data into SchoolCity for school staff to view. LEA-1-ES2 and LEA-1-MS waited to examine the data when the district made them available in SchoolCity. However, LEA-1-ES1 was very data-driven and began looking at results as they came in directly from the Test Operations Management System (TOMS). This information was shared with teachers at the school prior to when the information was available in SchoolCity. The POC knew data for the 2017–18 school year available in August were preliminary, but at this time they looked at the data and used existing cut points to identify the students who met or exceeded the mathematics and ELA standards. They determined the performance of LEA-1-ES1 overall and at each grade-level. The district eventually provided the same information. Speaking of the 2018–19 summative assessment, the LEA POC noted frustration that the summative assessment preliminary data were not available as soon after testing as they expected. They had been led to believe preliminary scores would be available immediately after testing. However, there was some wait time before ETS and CDE had results available. The more immediately results can be shared, the more useful they are to educators. Teachers at LEA-1-MS indicated some teachers accessed and used results, and others did not. One teacher mentioned they struggled with SchoolCity, which created a barrier to accessing results.

The LEA-1 POC noted that each school received data at the overall and claims level in SchoolCity. They're able to drill down to school-, grade-, and teacher-level. In 2018–19 they were able to pinpoint strengths and weaknesses. School POCs and teachers reported varied utility of the SchoolCity information. Some felt that SchoolCity was not very user-friendly for summative reporting. Teachers from LEA-1 schools noted they had access to data on SchoolCity; teachers at the elementary schools were also provided printed reports. Multiple elementary teachers expressed they had reviewed individual student reports to examine student performance and annual growth.

School and district-staff across LEA-1 did not uniformly agree on the utility of summative assessment results for informing instruction. LEA-level staff found the data first became useful as they began to see score improvement in 2018–19, because in previous years the LEA was at the bottom for all areas. The LEA-1-MS POC noted a school-level decision to focus more heavily on mathematics in 2018–19 than in previous years after learning that mathematics scores declined on the 2017–18 summative assessment. A

teacher from LEA-1-ES2 thought the information was somewhat useful, however found it tough to apply the information in the classroom. Some teachers at LEA-1-ES2 had a tough time recalling at the time of the focus group what information was shared.

Student Preparedness

Teachers and school-leaders discussed having prepared their students for the summative assessment by exposing them to online tests. Particularly, teachers and school POCs indicated use of CAASPP IABs, practice and training tests, as well as online assessments created in SchoolCity. The elementary schools noted use of Chromebooks from an early grade to increase familiarity for young students. During focus groups, students from all three schools stated they had enough practice with technology before the test, and many felt the questions were of similar format to what they had seen before.

Student focus group data from all three schools provided information on the student perspective of the summative assessment. Students at LEA-1 expressed that they felt they did the best that they could on the ELA and mathematics summative assessment. They reported that they focused and read the questions carefully.

Through focus group responses HumRRO found that students who found the ELA test easy indicated it was because they covered the material in class. However, some students at all three LEA-1 study schools indicated that the test was harder than anticipated. A couple students indicated the vocabulary was difficult. One middle school student felt that the ELA test was difficult because they had to write an essay, and one felt the questions required in-depth analysis of the text.

Students also had mixed opinions about the difficulty of the mathematics summative assessment. At all three study schools, some students found the mathematics test easy and some thought it was difficult. At LEA-1-ES1 some students indicated they didn't understand some problems, including fraction problems and word problems. Two students at LEA-1-ES2 thought the performance task was hard. One LEA-1-MS student thought the questions attempted to trick them – for example, asking for a unit of measurement you had to convert from another. A middle school student thought the performance task was tough because they had to refer back to earlier parts as they responded.

In conversations with school POCs about the student focus groups, they noted that many of their students were confident in their performance on the assessment; although the two elementary POCs weren't certain the confidence was warranted.

A few student focus group respondents at LEA-1-ES1 and LEA-1-MS provided information on their use of summative assessment tools. Some students indicated finding the highlighter tool useful on the summative. One middle school student indicated they often had trouble with focus, and the highlighter helped keep them organized and on task. This student also appreciated the ability to cross out text and having access to the calculator for math. One middle school student found the glossary useful, though another did not find it useful.

During the end-of-year focus group, the school POC from LEA-1-MS noted students struggle with reading comprehension of the questions on the summative, so more work may be needed to prepare students for that aspect of the assessment.

Interim Assessments

General Interim Assessment Information

LEA-1 did not mandate the use of interim assessments for the 2018–19 school year; they found schools did not respond well to mandates in the past. Instead, they focused on highlighting the benefits of IABs, and encouraged schools to use them. School POCs and teachers at all three schools indicated interim assessment decisions were generally made in PLCs or other teacher group collaboration. The LEA-1 CAASPP Coordinator provided HumRRO staff with access to teachers across the LEA. This differed from some districts where school CAASPP coordinators were responsible for this action.

Table C3 presents the total number of IAB tests taken by students in the three LEA-1 study schools during the 2018–19 school year, and the number of IAB tests taken by students in all LEA-1 schools. Counts of tests include those for students who took the same test multiple times. The table also indicates how many enrolled students in the LEA and each school are eligible to take the CAASPP Summative Assessments. The table also indicates how many of the total IABs were in each domain (ELA or mathematics) and how many were given in a standardized manner versus a nonstandardized manner. Only one Interim Comprehensive Assessment (ICA) was administered in the entire LEA; it was an ICA in mathematics and was administered in a nonstandardized manner.

Table C3. Number of Smarter Balanced IABs Taken by LEA-1 Students

LEA or School	CAASPP Eligible Students	Total # IABs ELA and Math	# ELA IABs	# Math IABs	# Standardized IABs ELA and Math	# Non-Standardized IABs ELA and Math
LEA-1	10,764	26,192	8,870	17,322	15,240	10,952
LEA-1-ES1	566	3,417	1,754	1,663	836	2,581
LEA-1-ES2	169	344	319	25	344	-
LEA-1-MS	812	1,848	58	1,790	1,552	296

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-1 had 10,764 students eligible for the CAASPP summative assessments in 2018–19. LEA-1 gave 26,192 total IABs (count of tests given). Of these, 8,870 tests were for ELA and 17,322 tests were for math. Of the total IABs, 15,240 were given in a standardized manner and 10,952 in a nonstandardized manner (across ELA and math).*

LEA-1 administered the IABs in various ways, both standardized and nonstandardized. However, some teachers at LEA-1-MS were unaware that IABs could be administered in a nonstandardized manner, indicating they were unaware they could administer the assessments as a full class or have students work in pairs. The LEA-level TOSAs are available to provide support for using IABs. They have provided professional development in the district and were present at the school when extra support was requested.

There were a few LEA-1-ES1 and LEA-1-MS teachers who chose not to use IABs for various reasons. One teacher at LEA-1-ES1 indicated their grade group had chosen to use another formative assessment that better met their needs. Multiple LEA-1-MS teachers indicated they used other formative assessments and adding IABs would have been too much testing for their students. Two of these teachers explained that their alternate formative assessments better matched their instruction and one felt the reporting in formative assessments generated by SchoolCity was more user friendly compared to the CAASPP IABs.

Table C4 presents the count of testing opportunities (i.e., test sessions) there were for specific ELA and mathematics IABs, by grade level, across the three LEA-1 study schools during the 2018–19 school year. Frequency of administration of an IAB for some grades includes more than one school. IABs that require hand scoring are noted in the table with an asterisk (*). Grade five administered the most ELA IABs (26), grade three administered the most mathematics IABs (20), and grade five administered the most overall number of IABs (45). The most frequently administered ELA IAB was Read Informational Texts (9 times in grade five). The most frequently administered mathematics IABs were Number and Operations in Base Ten (7 times in grade five) and Operations and Algebraic Thinking (7 times in grade 3). In the table, NA indicates the IAB is not available at that grade.

Table C4. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-1, by Domain and Grade

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
ELA	Brief Writes*	3	0	1	0	0	0
ELA	Editing	1	0	6	6	0	0
ELA	Language and Vocabulary Use	1	1	2	0	0	0
ELA	Listen/Interpret	1	1	1	3	1	1
ELA	Read Informational Texts*	4	4	9	1	3	0
ELA	Read Literary Texts*	4	4	4	2	0	0
ELA	Research	1	1	1	1	1	1
ELA	Revision	1	0	1	3	0	0
ELA	Performance Task*	1	1	1	0	0	0
ELA	**SUBTOTAL, all ELA IABs**	**17**	**12**	**26**	**16**	**5**	**2**

Table C4. (cont.)

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Math	Measurement and Data	6	0	1	NA	NA	NA
Math	Number and Operations-Fractions	1	1	3	NA	NA	NA
Math	Number and Operations in Base Ten	2	0	7	NA	NA	NA
Math	Operations and Algebraic Thinking	7	1	6	NA	NA	NA
Math	Geometry	3	0	1	1	1	1
Math	Expressions and Equations	NA	NA	NA	1	3	NA
Math	Ratios and Proportional Relationships	NA	NA	NA	3	1	NA
Math	The Number System	NA	NA	NA	3	3	0
Math	Statistics and Probability	NA	NA	NA	1	1	NA
Math	Expressions and Equations I	NA	NA	NA	NA	NA	1
Math	Expressions and Equations II	NA	NA	NA	NA	NA	1
Math	Functions	NA	NA	NA	NA	NA	6
Math	Performance Task*	1	1	1	0	0	0
Math	**SUBTOTAL, all Math IABs**	**20**	**3**	**19**	**9**	**9**	**9**
BOTH	**TOTAL IABs, ELA and Math**	**37**	**15**	**45**	**25**	**14**	**11**

IABs that require hand scoring are noted in the table with an asterisk (*).

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at each grade across the LEA-1 schools in the study. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-1 there were 3 testing opportunities for Brief Writes at grade 3, none at grade 4, and 1 at grade 5. There were no Brief Writes testing opportunities at grades 6, 7, or 8.*

When IABs were new, LEA-1 required hand scoring across the district. District staff were trained and came back and trained all teachers. However, the LEA has since made IABs voluntary, including which are selected. LEA-1 staff indicated that though some teachers may avoid hand scoring, those who have been trained often find it very doable. Teachers and school leaders at the two elementary schools indicated selecting IABs with hand scoring intentionally; teachers in focus groups noted the higher level of thinking required. One teacher at LEA-1-ES2 noted an IAB requiring hand scoring was selected because of its coverage of the essential standards. All teachers at the middle school who provided input to HumRRO's study indicated not using IABs requiring hand scoring in 2018–19.

LEA-1-MS discussed the use of designated supports during IAB administration. Teachers were unclear about what designated supports could be made available to their students, though they understood that accommodations were only available for (a)

students with disabilities who had an Individualized Education Program (IEP) or (b) students with a 504 Plan. The CDE notes that educators can make recommendations for designated support usage, then an LEA CAASPP Coordinator or Site Coordinator is responsible for turning on the recommended designated supports.

Purpose of Interim Assessments

All three LEA-1 study schools included teachers and/or school leaders who used interim assessments to prepare students for the summative assessment. One teacher at LEA-1-ES1 noted they began administering IABs in March as preparation for summative testing. The LEA-1-ES1 POC echoed this purpose and indicated that at mid-year the teachers began to focus on getting students familiar with the IABs so the summative assessment "wasn't quite so scary." A teacher at LEA-1-MS noted a mathematics IAB was given to familiarize students with all aspects of the technology involved in taking the assessment, including universal tools and the calculator. Similarly, some teachers noted selecting IABs, particularly IABs that required hand scoring, because they require higher-level thinking similar to what is required on the summative.

Teachers from LEA-1-ES1 indicated use of IABs to inform instruction. Particularly, the school POC noted teacher groups discussed IAB results during PLC time and identified gaps in student knowledge. They used the information to target instruction. One teacher at this school indicated an ELA IAB was selected specifically to gather student data on their language arts progression. They found the IAB included the content they were looking to assess. This school has embraced IABs and uses them frequently for mathematics and ELA. The LEA POC noted staff at LEA-1-ES2 elementary school selected an IAB requiring hand scoring because it would provide information on student ability on essential standards. A teacher at LEA-1-ES2 echoed this intent, indicating they discovered students needed additional practice in an area after hand scoring an IAB. Similarly, the LEA-1-ES2 school POC noted that though they are early in the process, their intent is to use IABs to inform instruction. One LEA-1-MS teacher noted one goal of administrating a mathematics IAB was to monitor the students' progress in mastering the standards covered by the IAB. This same teacher indicated plans for a mathematics IAB to become a common assessment across all middle schools in the district in order to share successes of the students' progress. One LEA-1-ES1 teacher expressed IAB results were used to help guide development of Distributive Practice.

Interim Assessment Reporting System

Prior to IAB use, the LEA POC completed rostering for all 20 schools in the district and for every teacher as necessary for obtaining IAB results. Approximately 600 teachers used IABs, and some have multiple classes. This took some time and meant some teachers who wanted to administer IABs early (September 1) were not able to do so. However, the LEA POC felt the time and effort was worth it. One school CAASPP coordinator expressed interest in having this control at the school-level – which is a possibility if the LEA decides to provide its school sites with direct access to the IA Reporting System.

At the school level, school POCs and teachers at LEA-1 most commonly indicated viewing results in the IA Reporting System. There was consistency across the three schools regarding the type of information viewed. Teachers indicated they examined scores for their class overall and for individual students. Some teachers indicated digging further into data, noting they reviewed the specific targets students did well on and those they struggled with. They also examined results for each item to identify difficult questions.

Digital Library

LEA-level staff indicated some level of DL use in their schools. For example, a LEA staff member who conducted professional development in schools indicated that since educators have been using IABs they have had increased exposure to the Digital Library through the Playlist. She noted educator excitement as they found relevant resources to meet their students' needs following IAB administration. Additionally, she noted satisfaction with the ability to tag potentially useful resources for later use. The LEA POC noted that the DL is only barely being used across the LEA at this point because the focus has been on creating strong PLCs and systems change; however, with the expected increased use of IABs, she expects use of the DL resource will increase as well.

Although LEA staff and some school leaders expressed a generally positive perspective of the Digital Library and its utility, there was limited use of the resources in 2018–19 by educators across the three LEA-1 study schools. A LEA-1-ES1 teacher indicated that another teacher who had participated in the CAASPP Summer Institute shared information about parts of the digital library; however, it still feels a bit foreign and teachers need more time to delve in. An LEA-1-MS teacher stated they had a wealth of other excellent resources, so they had not used the DL. Particularly, LEA-1 staff indicated having access to many resources through their ELA curriculum, StudySync, and through Achieve 3000. However, if they came across an especially strong lesson related to a standard their PLC team is working on, they would use it.

LEA-2 Findings

The LEA-2 participation included one elementary school, one middle school, and one high school. Table C5 summarizes the qualitative data available for analysis. As shown, the elementary school was a consistent contributor of information to this study, with strong collaboration from the LEA throughout the duration of the project. The middle school and high school participated in site visits, but none of the other sources of data collection except for middle school participation in the first monthly polling.

Table C5. Summary of Data Sources for LEA-2

Data Source	Participants/Description
Site Visit Educator Focus Group	ES – 9 teachers across three focus groups: Kindergarten (1), grade one (1), grade two (1), grade three (2), grade four (2), and grade five (2) MS – 6 teachers across two focus groups: grade six (4), Instructional Support (RSP) Language Arts/Math (1), grade seven English/Computer Literacy/Yearbook (1), HS – 6 teachers across two focus groups: ELA (3), math (3)
Site Visit Leader Interview	ES – School POC (Principal) and Vice Principal; 1 interview MS – School POC (Principal); 1 interview HS – School POC (Principal), Assistant Principal, and LEA Coordinator; 1 interview LEA – LEA POC (LEA Coordinator); 1 interview
Monthly Polling	ES – December (POC); January (5 teachers); February (5 teachers); May (1 POC) MS – December (1 teacher) LEA – December (POC), January (POC), February (POC)
End-of-Year Virtual Focus Groups	ES – POC LEA – POC
Student Focus Groups	ES – 9 students
Documentation	• District calendar, school schedules, and course lists • Smarter Balanced claim and target information, and blueprints • Slides and handouts summarizing district and school summative assessment performance and growth, overall and by subgroups, and compared to similar districts

LEA and School Characteristics

LEA-2 is a small district in southern California. The district includes three elementary schools, one middle school, one high school, and one alternative high school. Due to continued growth in the district, a new elementary school is scheduled to open for the 2019–20 school year. Table C6 presents a summary of demographic and achievement characteristics for LEA-2 and its three participating schools. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest). As shown, about 46 percent of the students in the participating schools are from lower income families. Through site visit focus groups and interviews, HumRRO learned many students will be first-generation college students, and some will be the first in their family to graduate from high school. The home language of many students is Spanish. More than one-third of LEA-2-ES students and slightly more than one-quarter of middle school students in the district are classified as English learners.

Table C6. Demographic Characteristics of LEA-2 and Its Participating Schools

Variables	LEA-2	LEA-2-ES	LEA-2-MS	LEA-2-HS
Enrollment	4,270	866	979	1,249
% Socioeconomically Disadvantaged	49%	46%	48%	43%
% Students with Disabilities	8%	9%	11%	7%
% English Learners	26%	36%	26%	11%
% Reclassified Fluent English Proficient	17%	2%	16%	37%
% Met or Exceeded ELA State Standards	59%	59%	60%	71%
% Met or Exceeded Math State Standards	44%	48%	45%	31%

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA overall and for each of the LEA's schools in the study. The second column (from top to bottom) shows in LEA-2 there was a total enrollment of 4,270 in 2017–18. Of these students, 49% were socioeconomically disadvantaged, 8% were students with disabilities, 26% were English learners, and 17% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 59% of students met or exceeded ELA state standards and 44% of students met or exceeded math state standards.*

Between 7 and 11 percent of students in the participating schools are identified as students with disabilities. Instruction for these students is generally in traditional classrooms using an immersion approach with appropriate resources. HumRRO staff spoke with three elementary school teachers who serve on the Student Study Team (SST) committee to support struggling students. They meet as a team of teachers and administrators along with parents and develop a plan to support the student. If a student continues to struggle, they screen the student to see if they qualify for special services. LEA-2-ES has a classroom for students with severe disabilities.

ELA and mathematics achievement of the participating elementary school and middle school is similar to overall district achievement, with about 60 percent meeting or exceeding standards in ELA and less than half (48 and 45 percent, respectively) at least meeting mathematics standards. High school achievement is considerably higher than overall district achievement in ELA but much lower in math.

Teachers from LEA-2-ES who participated in the focus group had a wealth of teaching experience, ranging from 3 to 24 years with most in the teaching profession for 18 years or more. With relatively low turnover, teachers bring considerable experience of the community to their classrooms. The middle and high school teachers were slightly less experienced, with an average of 14 and 11 years of experience, respectively.

Professional Learning Communities and Professional Development

LEA-2 frequently used a train-the-trainer model for professional development. They typically invited an administrator and one or two teachers representing ELA or mathematics from each school site to attend professional development training provided by the County Office of Education. These individuals were expected to train others at their school. They were offered formative and summative assessment training as well as training on the DL. Training participants shared handouts and webinars with school staff. With a large population of English learners, LEA-2 benefitted from the recent addition of training in the use of English Language Proficiency Assessments for California (ELPAC). The County Office of Education expects to support more training, especially in mathematics since they recently hired a mathematics training specialist.

Initiatives and Resources

Teachers at all grade levels in LEA-2 indicated their curriculum included a wealth of resources to support instruction and learning. They described the amount of material as sufficient, with some teachers reporting the breadth of resources and supplemental material as overwhelming. The curriculum provides interim assessments aligned to the CCSS.

School POCs and teachers across LEA-2 discussed using SchoolCity to manage student assessment results in prior years. LEA-2 no longer subscribes to the service.

Technology

Access to computers varied across classrooms in LEA-2's participating schools. The internet broadband had been increased and updated to meet the demands of digital devices. LEA-2 purchased Google Chromebooks for teachers and schools. Each elementary teacher had Chromebooks to use in the classroom and to administer interim and summative assessments. Students in grades three, four, and five had their own earbuds and mice; they regularly used the Chromebooks. At LEA-2-MS, teachers did not have as many Chromebooks as they needed. Half of the grade 6 teachers had Chromebooks and the others used a technology cart or the computer lab. LEA-2-HS had carts of Chromebooks, but not enough for every teacher.

Use of CAASPP Components

LEA-2 was taking a gradual approach to integrating CAASPP interim assessments and data. The elementary school was further along than the middle school, which was further along than the high school. The administrators at LEA-2-ES had made CAASPP data a centerpiece of school improvement. At the beginning of the school year, teachers of students in grades three through five received proficiency level results from summative assessments for students they taught last year and students they will teach during the current school year (except for third grade students who have no prior year results). Teachers also received comparison data from the other elementary school in their district as well as summative scores from two other districts. Middle school teachers received overall scale scores, but no claims- and target- level data. High

school teachers had not been able to fully access CAASPP data. This stems from having had access only to data for students in their current classes and none from their former grade eleven students. Ninth grade teachers received data from eighth grade summative assessments. Teachers of eleventh graders had no prior year data for those students, so they were unable to review data to identify student strengths and weaknesses and adjust instruction for them.

The district previously used SchoolCity to assess and manage students. They no longer had access to SchoolCity in 2018–19, but their CCSS-aligned curriculum provided classroom assessments, including performance tasks. The curriculum offered myriad supplemental resources, thus most teachers reported having sufficient formative assessments within their curriculum materials. The district was committed to using the CAASPP interim assessments in conjunction with classroom and curriculum assessments. However, district and school administrators had not mandated CAASPP interim assessments. They expected teachers to incorporate at least some interim assessments into their instruction. At the elementary school, teachers met in grade-level groups to identify relevant interim assessments and provide a testing plan to the administration. Middle school and high school teachers were also encouraged to use the CAASPP interim assessments, but the process is not as formalized as in LEA-2-ES. In grades six through eleven in LEA-2, teachers typically made individual decisions to use interim assessments in their classrooms.

Summative Assessments

LEA-2 had secured internet bandwidth and Chromebooks to allow students the opportunity to take practice summative assessments. Students were familiar with the tools and format of the assessments as well as the expectations and rigor. Teachers in the elementary and middle schools received some summative assessment data and were learning to use it to inform their instruction and identify student gaps and need for additional instruction. LEA-2 administrators and staff are working towards increased availability and integration of CAASPP assessment data across the district, particularly in the middle and high schools.

Access to and Use of Summative Data

Elementary teachers use summative assessment results to identify areas of students' strengths and weaknesses. Staff at the elementary level use this information to examine their curriculum by grade level to inform instruction to support student learning and progress. For example, grade five teachers identified a need for additional resources to supplement the curriculum and provide more practice in writing. They also recognized a need to strengthen students' basic multiplication skills and understanding of geometry. Assessment content has led teachers to change the order in which they cover specific topics to ensure their students have learned grade-level content standards prior to the assessment (i.e., not going sequentially through a textbook).

Middle school teachers reported receiving summative assessment raw scores, but no data by claims. To access claim-level data, they could use paper copies of parent reports in parent conference folders. By focusing on claims and targets, middle school

ELA teachers identified listening and speaking as areas for improvement. Teachers at LEA-2-MS were encouraged by administrators to identify students who were close to meeting the standard, but teachers did not find that information useful as it did not indicate the specific skills students needed help with. Teachers asked for "actionable data" on "skill-based areas of need."

At LEA-2-HS, teachers were provided summative assessment results for students currently in their classes. Grade nine teachers were provided access to grade eight results for their current students. Grade eleven teachers cannot see student-level data of the students they taught last year once the students are promoted to grade 12. High school teachers have access to district-level data. Based on this, LEA-2-HS is focusing on improving performance in mathematics. ELA teachers at LEA-2-HS have not used summative assessment results to impact classroom instruction.

Student Preparedness

Elementary school students and their teachers generally reported being confident students performed well on the summative assessments. Most elementary students who participated in student focus groups said they were prepared for the rigor of the assessment as well as the computer format, including the types of questions and the tools. Students reported using the glossary and notes for writing. They also used the magnifier, highlighter, strike-through, line reader, and color contrast. School staff attribute student readiness, in part, to the curriculum and practice assessments aligned to the state assessments. Students referenced taking the IABs on Chromebooks as helpful in preparing them for the summative assessments.

Several students and teachers commented some questions were confusing. A teacher reported students knew the mathematics content but did not understand the questions. Some students found selected questions difficult and some content they had not been taught during the current school year, including mathematics topics such as measurement, ratios, angles and degrees, commutative property, and three-digit estimation. Students indicated some of the unfamiliar and less familiar content was taught during the previous year.

Middle school and high school teachers reported their students were comfortable in using technology. In the high school, teachers assigned online homework and exams to prepare students for the summative assessments.

Interim Assessments

General Interim Assessment Information

LEA-2 implemented IABs for the first time in 2017-2018. They did not mandate use of the interim assessments but expected and encouraged teachers to use them. Teachers were asked to report the IABS they used and when. Many teachers, particularly in the elementary schools, used them throughout the school year. Elementary school teachers decided by grade level how to use the IABs (i.e., mapping out which IABs to use and when to take them) to prepare their students for the content, rigor, and format of the

summative assessments. One teacher preferred using the shorter mathematics assessments because the ELA IABs took two hours of instructional time. In some cases, IABs were completed collaboratively as a class and the teacher modeled taking the assessment with student discussion. Teachers generally did not use accommodations with the IABs, especially when used in a group setting. Middle school teachers had just started using the interim assessments. Getting the high school to use the interim assessments had been challenging. Decisions to use interim assessments at the high school were made at the department and administration levels. The district goal was to administer interim assessments at least quarterly for 2018–19 and to implement ICAs in future years.

Table C7 presents the total number of IAB tests taken by students in the three LEA-2 study schools during the 2018–19 school year, and the number of IAB tests taken by students in all schools in the LEA. Counts of tests include those for students who took the same test multiple times. The table indicates how many enrolled students in the LEA and each school are eligible to take the CAASPP Summative Assessments. The table also indicates how many of the total IABs were in each domain (ELA or mathematics) and how many were administered in a standardized manner versus a nonstandardized manner. LEA-2 administered the majority of IABs in nonstandardized manner. No IABs were administered in high school, and no ICAs were administered across the entire LEA.

Table C7. Number of Smarter Balanced IABs Taken by LEA-2 Students

	CAASPP Eligible Students	Total IABs	ELA IABs	Math IABs	Standardized IABs	Non-Standardized IABs
LEA-2	2,188	8,194	4,728	3,466	1,257	6,937
LEA-2-ES	418	2,675	964	1,711	474	2,201
LEA-2-MS	974	4,137	3,122	1,015	432	3,705
LEA-2-HS	272	-	-	-	-	-

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-2 had 2,188 students eligible for the CAASPP summative assessments in 2018–19. LEA-2 gave 8,194 total IABs (count of tests given). Of these, 4,728 tests were for ELA and 3,466 tests were for math. Of the total IABs, 1,257 were given in a standardized manner and 6,937 in a nonstandardized manner (across ELA and math).*

Table C8 presents the count of testing opportunities (i.e., test sessions) there were for specific ELA and mathematics IABs, by grade level, across the three LEA-2 study schools during the 2018–19 school year. Frequency of administration of an IAB for some grades includes more than one school. IABs that require hand scoring are noted in the table with an asterisk (*). Grade seven administered the most ELA IABs (35), grade three administered the most mathematics IABs (21), and grade seven administered the most overall IABs (46). The most frequently administered ELA IAB was Read Informational Texts (10 times in grade seven). The most frequently administered mathematics IAB was the Performance Task (6 times in grade three). In the table, NA indicates the IAB is not available at that grade.

Table C8. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-2, by Domain and Grade

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	HS
ELA	Brief Writes*	3	1	1	1	3	1	1
ELA	Editing	1	1	0	1	3	0	0
ELA	Language and Vocabulary Use	3	3	3	6	3	0	1
ELA	Listen/Interpret	3	1	0	1	3	1	1
ELA	Read Informational Texts*	3	1	1	3	10	1	1
ELA	Read Literary Texts*	3	3	1	3	6	1	1
ELA	Research	1	1	0	0	3	0	1
ELA	Revision	1	0	0	1	3	0	1
ELA	Edit/Revise	NA	NA	NA	NA	NA	1	NA
ELA	Performance Task*	3	1	0	1	1	0	0
ELA	**SUBTOTAL all ELA IABs**	**21**	**12**	**6**	**17**	**35**	**5**	**7**
Math	Measurement and Data	3	1	1	NA	NA	NA	NA
Math	Number and Operations - Fractions	3	3	1	NA	NA	NA	NA
Math	Number and Operations in Base Ten	3	1	3	NA	NA	NA	NA
Math	Operations and Algebraic Thinking	3	1	1	NA	NA	NA	NA
Math	Geometry	3	1	1	0	1	1	NA
Math	Expressions and Equations	NA	NA	NA	1	3	NA	NA
Math	Ratios and Proportional Relationships	NA	NA	NA	3	3	NA	NA
Math	The Number System	NA	NA	NA	1	3	1	NA
Math	Functions	NA	NA	NA	NA	NA	1	NA
Math	Algebra and Functions I	NA	NA	NA	NA	NA	NA	1
Math	Algebra and Functions II	NA	NA	NA	NA	NA	NA	1
Math	Number and Quantity	NA	NA	NA	NA	NA	NA	1
Math	Geometry Congruence	NA	NA	NA	NA	NA	NA	1
Math	Geometry Measurement and Modeling	NA	NA	NA	NA	NA	NA	1
Math	Performance Task*	6	1	1	1	1	0	0
Math	**SUBTOTAL all Math IABs**	**21**	**8**	**8**	**6**	**11**	**3**	**5**
BOTH	**TOTAL IABs, ELA and Math**	**42**	**20**	**14**	**23**	**46**	**8**	**12**

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at each grade across the LEA-2 schools in the study. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-1 there were 3 testing opportunities for Brief Writes at grades 3 and 7, one each at grades 4, 5, 6, 8, and high school.*

Purpose of Interim Assessment Use

Teachers used the IABs to model and teach students how to take the assessment. For third graders, the IABs provided typing practice. Some teachers incorporated the IABs as a teaching tool, using assessments as practice in areas where students struggle. Teachers were just beginning to use standardized IABs to measure student progress, identify strengths and weaknesses, gaps in learning and understanding, and inform instructional decisions. A teacher reported reviewing hand-scored items and their rubrics with students so they will have a better understanding of what is expected on the summative assessment. One elementary teacher commented "the IAB is a great tool to ensure we are teaching standards and providing our students with opportunities to be successful." Teachers also shared individual student scores with students to provide feedback.

A few high school staff used interim assessments in past school years to expose students to the types of questions and expected responses on the summative assessments, particularly mathematics items with multiple steps. LEA-2-HS conducted voluntary Saturday CAASPP practice sessions with students. One teacher reported taking all IABs to become familiar with the assessment content and used it to inform classroom instruction to ensure students were prepared for the breadth of material on the summative assessment. This teacher may not have realized they could have used the Teacher Viewing System to view all the IABs without taking the IABs. Teachers expressed concern for needing time and access to computers for students to practice typing and to prepare them to write on a computer, which taps different cognitive processes than writing on paper. During the site visit, LEA-2-HS teachers noted they likely would not be using IABs during 2018–19.

Interim Assessment Reporting System

Teachers and administrators across the district spoke about using School Island, a learning management system, in prior years to store and access student performance data. They no longer subscribe to this system and describe the current reports as difficult to access and use because they often must extract information for students in their class out of all students in a grade or in the school. High school teachers are only able to see score reports for their current students. Grade nine teachers have access to students' grade 8 scores. However, grade eleven teachers are not able to see results from last year to identify student strengths and weaknesses to inform their instruction, because those students are no longer in their class. As a proxy, they use school-level data.

At LEA-2-ES, one teacher hand scored an interim assessment but was unable to access the results. Another teacher reported being unsure how to hand score. A teacher noted hand scoring is subjective and discrepant even when using the provided rubrics. This teacher suggested collaborative training to make scoring more consistent within each grade. Some staff received training on scoring but requested additional training for hand scoring interim assessments.

High school teachers say they do not have time to hand score interim assessments, and cited this as a barrier for them to implement IABs. They described difficulties using interim assessments in a standardized manner because the state will not grade the multiple-choice questions unless the performance task is scored by the teacher. In addition, when teachers score the performance task and other questions are machine scored, teachers believed they would not be able to print reports. However, this may have been a misunderstanding of the system, as teachers would be able to print a report once the PT is completely scored and the IA submitted for scoring. The high school teachers describe the interim assessment reporting system as not useful.

Digital Library

Most teachers reported they did not use the DL. Many never logged into the Digital Library. Those who accessed it described it as not user friendly, stating it was difficult to navigate. Staff acknowledge they did not receive enough professional development to use it. They used other resources, primarily their adopted curriculum materials which are extensive and "overwhelming." Most teachers agreed they do not have enough time to review and locate relevant material in the DL. These teachers may not have been aware that they could access relevant materials directly from the Instructional Resources button in the IA Reporting System. However, several teachers said they did use the library as a useful resource. One elementary teacher has taken ideas, but not complete lessons, from the DL. An elementary teacher used the DL, accessed through the Playlists related to IA results, for resources to reteach concepts for students near or below standard. Some high school teachers reported finding limited content for their classes when the DL first opened and did not access it again.

LEA-3 Findings

Table C9 summarizes the LEA-3 qualitative data. As shown, LEA-3 participated in data gathering events throughout the school year with site visits at the beginning, a monthly polling in the middle, and end-of-year focus group participation. LEA-3, however, did not participate in the student focus groups or in most of the monthly polling events.

Table C9. Summary of Data Sources for LEA-3

Data Source	Participants/Description
Site Visit Focus Group	15 teachers across two focus groups – ELA (7) and math (8)
Site Visit Interview	School POC (CAASPP Coordinator) and Principal
Monthly Polling	January (11 teachers)
End-of-Year Virtual Focus Groups	POC (CAASPP Coordinator)
Documentation	Achieve3000_WWC Intervention Report

LEA and School Characteristics

LEA-3 is a direct-funded charter high school in southern California. Table C10 presents a summary of demographic and achievement characteristics for students in the school. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest). As shown, the majority of students come from socioeconomically disadvantaged homes. Relatively small proportions of the student population are students with disabilities or English learners, with teachers saying that typically one to five such students are enrolled in a given class. Teachers report about 20% of students are reclassified ELs. The school mainstreams most special education and EL students with the expectation that they do what other students do. LEA-3 students performed very well academically, with 83 percent meeting or exceeding the grade eleven standard for ELA and 73 percent doing so for mathematics.

Table C10. Demographic Characteristics of LEA-3 (Charter High School)

Variables	LEA-3
Enrollment	2,465
% Socioeconomically Disadvantaged	60%
% Students with Disabilities	9%
% English Learners	7%
% Reclassified Fluent English Proficient[1]	23%
% Met or Exceeded ELA State Standards	83%
% Met or Exceeded Math State Standards	73%

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA charter overall. The second column (from top to bottom) shows in LEA-3 there was a total enrollment of 2,465 in 2017–18. Of these students, 60% were socioeconomically disadvantaged, 9% were students with disabilities, 7% were English learners, and 23% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 83% of students met or exceeded ELA state standards and 73% of students met or exceeded math state standards.*

The school sets high academic standards for its students. The school uses A-G graduation requirements, the series of high school courses required to be eligible for admission to California universities, so every student is at the college prep level. Teachers described students as having a strong desire to represent their school in testing. They outperformed other local schools in testing results.

Enrollment in this charter school was not related to ability, but it did have geographical priorities. Students within the charter school boundaries were given first priority, followed by students living within the local school district, followed by students outside those boundaries. While most students joined the charter their freshman year, a fair number of students also entered as sophomores or juniors.

LEA-3 had implemented a unique class scheduling approach to provide more remediation during the regular school day. Each semester, students had four blocks for which most students selected three classes. The additional class period was used as an academic support period to help students master the material covered in their classes. Students in good standing could instead enroll in elective courses. Additionally, the open period was used in grade nine for a course that introduced students to the school culture and invited them to start thinking about and planning for their future.

Depending on the grade, some students are in ELA or mathematics courses only one of the two semesters. The Algebra I and II classes are year-long courses, whereas other ELA and mathematics courses are only one semester. This meant some grade eleven students were not enrolled in an ELA or mathematics course during the semester when the 2018 summative assessment was administered, so these students would not have had instruction relevant to the assessment for a few months before the testing window.

Teachers at LEA-3 had a range of experience. Many teachers who participated in data gathering events had worked ten years or more as teachers with much of that experience occurring at LEA-3. In general, teacher retention was high. Numerous teachers also taught courses at local colleges and universities. Classroom teachers could take on special roles in technology implementation and other special projects. Additionally, teachers play a meaningful role in school decision-making activities.

Professional Learning Communities and Professional Development

LEA-3 teachers were organized into PLCs at the course level. These groups were responsible for activities such as assessing summative results, determining when interim assessments were given, and creating end-of-unit assessments and associated practice tests.

LEA-3 employed a grade-level team (GLT) model for student support. The GLT was composed of a principal, counselor, and academic advisor who remained attached to a particular cohort of students throughout their high school experience. The GLT met weekly to discuss how to support students. They also interacted with course-level PLCs to monitor student academic progress.

Teachers reported attending training on the CCSS. They were able to receive funding for professional development opportunities oriented around CCSS teaching. Additionally, the CAASPP coordinator met with PLCs to train them on data analysis and data reporting systems.

Initiatives and Resources

LEA-3 implemented a number of resources to prepare students for the CAASPP summative assessments. Actively Learn and Achieve 3000 were resources the high school used to provide students with experience interfacing with reading texts online and returning to text to answer questions.

Technology

LEA-3 had a one-to-one computer program and a number of tech-dedicated staff. Teachers and administrators did not feel technology was a barrier to using CAASPP resources for them or their students.

Use of CAASPP Components

LEA-3 used CAASPP assessment data for particular uses. The summative assessment is used as an evaluation tool. Teachers described how one year's poor testing results led the staff to evaluate their chosen curriculum and learning activities, which led them to identify new resources. Since then, scores had improved, and summative results had not been examined to the same extent. Interim assessments were primarily used to introduce students to the CAASPP assessment interface. In 2018–19 only mathematics interim assessments were used. Teachers reported discomfort accessing summative and interim assessment data independently.

LEA-3 regarded CAASPP formative and interim resources as one among many possibilities in preparing their students for the summative assessments. They had recently adopted College Board's Springboard curriculum because the curriculum's "embedded assessments" closely align with the format of performance tasks in CAASPP summative assessments. These embedded assessments were used every couple of weeks as part of unit tests. For grade nine and ten students, for which IABs were not specifically available, NWEA tests were used to monitor student progress. Additionally, PARCC practice tests were used for practice and class warmups because they did not require log on for access and teachers. Teachers also developed and administered their own "Smarter Balanced-like" questions and timed writings and administered those using their Canvas system. Other teachers researched Smarter Balanced terms and had discussions with students on what they meant.

Summative Assessments

Access to and Use of Summative Data

Summative assessment results from 2017–18 were provided to teachers at the beginning of the year in an all-teacher meeting. These results documented the overall achievement level score and year-to-year growth across the whole school and by different demographic groups. Additionally, teachers were provided comparison data to examine how their school did compared to other schools in the state. Teachers reported the claim- and target-level results were not presented at this all-teacher meeting. Teachers did not use these data to inform their instructional decisions. They came away from this meeting desiring more granular data about their domain (ELA or math) and data at the individual student level. They believe such data would help teachers understand how the previous year's innovations helped student achievement as well as guide decision making in the classroom. Teachers at LEA-3 did not know how to access data beyond what they were provided.

Later in the year, the CAASPP Coordinator met with each of the PLCs at the school to present and discuss summative assessment results. The Coordinator extracted the individual student results and mapped those to teachers to create reports. These data were intended to help the PLCs understand the CAASPP claim- and target-level results to guide goal setting and instructional planning. Teachers were shown how to use ORS. The results from the ORS reports resulted in discussions about how the teachers should focus on the different elements of the curriculum. Teachers also intended to use the summative assessment data to evaluate how curriculum implemented in the previous year influenced results. In a separate data gathering event, a HumRRO researcher demonstrated the use of ORS to teachers, who recognized the interface and remembered receiving training by their CAASPP coordinator in how to use the system. However, teachers had not had adequate time to go in and use the system to this point, and thus were not yet comfortable with it. Teachers were not familiar with the computer adaptive scoring process and this made it difficult to interpret scores.

Student Preparedness

As noted above, the LEA-3 teachers made use of numerous resources with the intent of helping students prepare for the summative assessment. The Springboard curriculum was chosen for the similarity of the embedded tasks to Smarter Balanced performance tasks in terms of format and rigor. Similarly, the Achieve software was selected to provide students experience reading texts on a screen and navigating such an interface in an assessment. Teachers at the school ensured students were exposed to a variety of text genres and received timed writing prompts and other questions that mirror the format of the summative assessment. The IABs and CAASPP practice tests were also used so that students could "experience the quirkiness of the [summative] exam." In the end-of-year focus group, the high school's CAASPP Coordinator indicated he believed students felt pretty good about the experience.

Interim Assessments

General Interim Assessment Information

Administrators strongly encouraged the use of IABs hoping teachers would see their benefits and adopt them as a resource. PLCs are empowered to decide among themselves which IABs are used for a given course and when they are administered. The ELA teachers chose not to use IABs in 2018–19. The mathematics teachers had a goal to administer an IAB that aligned with the given unit at the end of each quarter. Individual teachers could choose to administer more IABs, and some did; these teachers administered five different math IABs in the few weeks before the summative assessment. IABs are primarily used with grade eleven students.

Table C11 presents the total number of IAB tests taken by LEA-3 students during the 2018–19 school year. Counts of tests include those for students who took the same test more than one time. The table indicates how many enrolled students in the LEA are eligible to take the CAASPP summative assessments. The table also indicates how many of the total IABs were in each domain (ELA or mathematics) and how many were

given in a standardized manner versus a nonstandardized manner. No ICAs were administered in LEA-3.

Table C11. Number of Smarter Balanced IABs Taken by LEA-3 Students

	CAASPP Eligible Students	Total IABs	ELA IABs	Math IABs	Standardized IABs	Non-Standardized IABs
LEA-3	564	2,407	-	2,407	1,031	1,376

Explanation of table contents: *Row 1 shows LEA-3 had 564 students eligible for the CAASPP summative assessments in 2018–19. LEA-1 gave 2,407 total IABs (count of tests given). Of these, all 2,407 tests were for math. Of the total IABs, 1,031 were given in a standardized manner and 1,376 in a nonstandardized manner.*

Table C12 presents the count of testing opportunities (i.e., test sessions) there were for specific mathematics IABs at LEA-3 during the 2018–19 school year. LEA-3 did not administer any ELA IABs or any mathematics IABs that required hand scoring. The most frequently given mathematics IAB was Algebra and Functions I (6 times).

Table C12. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-3

Domain	IAB Name	HS
Math	Statistics and Probability	3
Math	Algebra and Functions I	6
Math	Algebra and Functions II	3
Math	Geometry and Right Triangle Trigonometry	3
Math	Interpreting Functions	3
Math	Number and Quantity	1
Math	Seeing Structure in Expressions/Polynomial Expressions	1
Math	**TOTAL Mathematics IABs**	**20**

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at LEA-3. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-3 there were 3 testing opportunities for Statistics and Probability.*

Through teacher focus groups HumRRO learned that hand scoring IABs caused frustration among some teachers. Teachers were unaware in advance that some IABs required hand scoring, and that results from IABs requiring hand scoring could not be examined unless the hand scoring section had been completed.

IABs were administered primarily in a nonstandardized way where students worked together and classes discussed the format. The emphasis was on students' experiencing the assessment, not on using or evaluating their results. Teachers

discussed challenges they encountered with the IABS. First, teachers and administrators suggested that the California Longitudinal Pupil Achievement Data System (CALPADS) needs to be updated sooner by their LEA to accommodate IAB usage earlier in the school year. Second, some teachers felt the Smarter Balanced summative assessment released items available through the CAASPP website differed from the IAB item types their students encountered. They hoped to see more consistency to ensure the IABs adequately prepared students for the summative assessments.

Purpose of Interim Assessment Use

A primary purpose for LEA-3 teachers' use of IABs was to familiarize their students with the summative assessment format. Teachers discussed item types while working through IABs as a class to help students understand what each item type required for an answer. One teacher described walking through IAB question types and explaining to students what was expected to receive credit for responding. Others described the value of student exposure to different tools within the interface (e.g., calculator). Even the experience of signing into the system was deemed valuable as teachers believed it would reduce stress when students took the summative assessment. LEA-3 teachers believed that increased exposure to computer test taking and the specific item types seen in the IABs provided great value.

A number of teachers suggested the IABs helped them assess how well their curriculum helped prepare students for the summative assessments. In some cases, teachers found that the IABs evidenced alignment of the curriculum with the assessment. One teacher described that in this way, "if anything, [the IABs] told us to keep doing what we were doing." For others, the IABs underscored areas that had not been focused on in the curriculum. This proved helpful as teachers could then use review periods to develop student mastery in these areas.

Other IAB uses were mentioned by individual teachers. One talked about using IAB results to determine what skills to focus on in mathematics class warmups. Others talked about using the IAB as a review at the end of the quarter.

Interim Assessment Reporting System

Teachers from LEA-3 accessed IAB reports through the IA Reporting System. The system was challenging for teachers in our study to use. Teachers at the site-visit focus groups found it difficult to log on and hard to interpret available reports. The CAASPP Coordinator requested a simpler interface for teachers to access interim results. Some teachers reported it was unclear why certain answers were marked correct or incorrect. They would have liked more information explaining how items are scored. However, the item rubrics have been available through the IA Reporting System since the 2017–18 school year – these teachers may have been unaware that this information was available.

Digital Library

One LEA-3 teacher had used the DL and had a positive experience. She recommended different material for Algebra. Beyond her experience, most teachers had not used it or had found it difficult to use. Teachers described difficulties in logging into and finding resources in the system. Teachers expressed frustration with the multiple different accounts required across the different CAASPP interfaces. Teachers also found the amount of resources overwhelming, which made it difficult to find relevant resources. Some indicated they had not accessed the DL since the previous school year. Teachers of AP courses stated their courses were not aligned to the Common Core, so the DL did not meet their needs. Other teachers felt they had sufficient resources from other sources and did not need the DL. One teacher indicated that across the CAASPP resources, she had been unable to find resources helpful in preparing for performance tasks.

LEA-4 Findings

HumRRO conducted the in-person site visit to LEA-4 in November 2018. However, in early December 2018, the LEA-4 POC requested to discontinue further participation due to the many district initiatives that were evolving and the lack of time for school staff to contribute to ongoing data collection. Though HumRRO offered to scale back the study activities to keep the LEA and its schools engaged, LEA-4 decided to conclude its role in the study. LEA-4's participation included one elementary school, one middle school (grades seven and eight), and one high school. See Table C13 for a summary of sources and participants in LEA-4's data collection activities.

Table C13. Summary of Data Sources for LEA-4

Data Source	Participants/Description
Site Visit Focus Group	ES – 3 teachers across three interviews: grade three (1), grade four (1), and grade six (1) MS – 9 teachers in one focus group: ELA (5) and math (4) HS –17 teachers in one focus group: ELA (5), math (6), and unknown ELA or math (6)*
Site Visit Interview	ES – School POC (Principal) and CAASPP Coordinator, 2 interviews MS – School POC (Principal), 1 interview HS – School POC (Principal) and CAASPP Coordinator, 1 interview LEA – LEA POC (CAASPP Coordinator) and Assistant Superintendent
Documentation	• 2017–18 IAB Student Achievement Handout • 2017–18 CAASPP Summative Results Presentation • 2017–18 CAASPP Summative Results Posters • 2018–19 Assessment Schedule

* Late arrivals to the focus group, subject area unknown.

LEA-4 and School Characteristics

LEA-4 is a small, rural district in an agricultural area of central California that includes five elementary schools, one middle school (grades seven and eight), and two high schools. Table C14 presents a summary of key demographic and achievement characteristics for students in the LEA and the three study schools. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest). As shown, a vast majority of students in this LEA (95%) are socioeconomically disadvantaged. Among the study schools, the English learner population is highest at LEA-4-ES (45%) and much lower at LEA-4-MS (18%) and LEA-4-HS (12%), where high percentages of ELs (53-60%) were reclassified as fluent English proficient. Students' academic achievement in the study schools was similar to that of the LEA overall; the majority of students did not meet or exceed the ELA or mathematics standards. Achievement at each study school varied by subject area; LEA-4-ES performed the best (31% met or exceeded ELA standards, 38% met or exceeded mathematics standards) while LEA-4-MS and LEA-4-HS had extremely low mathematics performance.

Table C14. Demographic Characteristics of LEA-4 and Its Participating Schools

Variables	LEA	LEA-4-ES	LEA-4-MS	LEA-4-HS
Enrollment	4,882	542	795	1,492
% Socioeconomically Disadvantaged	91%	88%	93%	91%
% Students with Disabilities	15%	10%	15%	12%
% English Learners	37%	45%	18%	12%
% Reclassified Fluent English Proficient	33%	19%	50%	57%
% Met or Exceeded ELA State Standards (2018)	30%	38%	18%	38%
% Met or Exceeded Math State Standards (2018)	19%	31%	10%	9%

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA overall and for each of the LEA's schools in the study. The second column (from top to bottom) shows in LEA-4 there was a total enrollment of 4,882 in 2017–18. Of these students, 91% were socioeconomically disadvantaged, 15% were students with disabilities, 37% were English learners, and 33% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 30% of students met or exceeded ELA state standards and 19% of students met or exceeded math state standards.*

Teachers highlighted the low socioeconomic status of the student population, with nearly all students receiving free-or-reduced lunch and a significant proportion lacking internet and computer access at home. At LEA-4-HS, the principal emphasized the isolation of the small-town community ("you have to have a car to get around") and

described students as behaviorally very compliant but not academically focused or rigorously engaged. He believed students were misled by teachers assigning higher grades than what may be deserved, which is not on par with their performance on AP and CAASPP tests.

In the last two years, LEA-4-ES experienced a high rate of teacher turnover and a high percentage of teachers were classified as probationary. In 2017–18, two-thirds of the teachers in each grade from three through six were on the pathway to a teaching credential, which meant many were simultaneously teaching, taking online courses, and studying for the California Subject Examinations for Teachers® (CSET®) and the Reading Instruction Competence Assessment® (RICA®). The principal, who had 13 years of administrator experience and was in his second year at LEA-4-ES, described the staff as "hard working with tons of potential." He noted they "have an extra challenge of the newness and lack of familiarity with the curriculum, grade level, testing, [and] working together as a team…They are learning pedagogy, methodology, curriculum, while they're sorting all this out." Similarly, at LEA-4-MS, nearly half the teachers and the principal were new to the school in 2017–18, and six of the 11 ELA and mathematics teachers were new in 2018–19 (including the entire grade seven ELA team). Of 35 teachers, 23 were credentialed but have only one or two years of teaching experience. At LEA-4-HS, the population of 65 teachers was stabilizing after two years of high turnover (loss of 30 teachers in 2016–17, 20 in 2017–18, and 12 in 2018–19).

Professional Learning Communities and Professional Development

To build teacher leadership and a shared approach to intervention and planning for student achievement, the LEA funded teams of principals and teachers in 2018 to attend PLC training sessions, and the LEA also supported weekly early release time for structured PLC meetings. At the three study schools, formal PLCs were described as being in their early phase.

At LEA-4-ES, 5 teachers were formally trained in the PLC process within the last year. Teachers described grade level teams of three teachers who meet weekly to collaborate and plan standards-based instruction to provide for consistent instruction and assessment across classrooms. For grade four, "Our team had a day to break down essential standards. We decided…which standards we were going to hit for math and which standards we were going to hit for ELA… what month we were going to hit them (approximately)." Teachers also noted their PLC time provided mentoring and support for the probationary teachers, "We have to be able to say when something isn't working and take advice from the rest of team" and "We need to put our egos at the door." Teachers described the physical proximity of their classrooms as further encouraging collaboration throughout the day.

At LEA-4-MS, 11 of 35 teachers were formally trained in the PLC process within the last year. PLCs met to collaborate in grade and subject level teams to identify essential standards, create common classroom assessments, and evaluate assessment scores (including district benchmarks) relative to teaching practices. One teacher commented that the time for this work, an hour and a half a week, isn't enough to "dive deep." With many new mathematics teachers, some PLC time is spent on learning the basics, such

as how to log on to systems. The principal stated he required each PLC to submit four completed products (following the PLC model from their training) twice a year: "Collaborative Team Products for a Unit of Study," "Collaborative Team SMART Goal Worksheet," "Collaborative Team Action Plan Worksheet," and "Collaborative Team Final Data Worksheet."

At LEA-4-HS, 13 of 65 teachers were formally trained in the PLC process within the last year, and these teachers served as PLC team leaders. PLCs are organized by subject area and grade-level, but sometimes all PLCs meet as a group. Teachers described the PLCs as educator driven, with team leads deciding what to work on each meeting. Examples of meeting topics included review of summative CAASPP scores, common lesson planning, and choosing common assessments (e.g., which IAB block to administer). PLC teams this year identified and documented the essential standards for each unit of study and determined how to distribute instructional time among the essential standards.

Initiatives and Resources

One initiative underway in the LEA is the identification of "essential standards" from the CCSS, versus use of the adopted curriculum including pacing guides, to guide focused instruction in ELA and math. Identification of the essential standards is a site-specific activity conducted by PLCs (grade level teams of teachers). During quarterly meetings of the LEA-4's Professional Learning Network (i.e., grade level teams assembled from teachers of multiple schools), teams discuss how to reach consistency across schools on the essential standards themselves, as well as how to assess the essential standards.

At all school levels, principals and teachers describe the approach for reteaching students during the school day has shifted from "pull-out" to "push-in." In prior years, the lowest achieving students in a class were sent from the classroom to meet with instructional assistants, who were less qualified than classroom teachers. As an outcome of PLC development, classroom teachers are working on interventions with their own low achieving students individually or in groups. At LEA-4-MS and LEA-4-HS, interventions to support struggling students also include before and after school tutoring and Saturday school.

The LEA-4-ES CAASPP coordinator was a Teacher on Special Assignment (TOSA) in the role of an EL resource teacher for the second year. When conducting English Learner Advisory Committee (ELAC) meetings, he found parents were surprised to learn students were struggling in reading and writing because the students demonstrate strong oral skills when speaking in English. Part of his role this year was to model lessons for teachers, such as having students work together in pairs, and observe teachers and provide feedback on how they deliver lessons. Similarly, LEA-4-MS received support from the district's academic coach (a TOSA) and the EL TOSA, each of whom supported multiple schools in LEA-4.

Technology

LEA-4-ES had two Chromebook carts with 30 computers, which could be checked out by teachers for classroom use, as well as a computer lab with about 40 Chromebooks. Students in grades three and up had opportunities to practice keyboarding throughout the year. In every classroom except kindergarten, each student had an iPad and, new this year, a keyboard for their iPad. In 2017–18, students took the Smarter Balanced summative assessments on Chromebooks, because there were not keyboards for the iPads and certain actions, like dragging and dropping with a mouse, are easier on laptops. A grade four teacher explained, "When we use Typing Club or other type apps, we use Chromebooks. We don't do a good job using the Chromebook carts like we should." A grade three teacher described the first time she introduced students to using laptops—for the purpose of practicing using Smarter Balanced assessment tools—as "like dropping a glass of marbles on the floor. It was difficult for students to focus. They were so overwhelmed…" In the first three months of the school year, the teacher had already conducted four practice sessions to familiarize students with the logging in process and use of tools such as highlighting and drag and drop.

LEA-4-MS and LEA-4-HS provided an iPad to every student for to use at home and school. Students take the Smarter Balanced interim and summative assessments on iPads, unless the tablet is about to run out of battery, in which case a Chromebook is used. The iPads had keyboards, but teachers noted many students at this age prefer to type with one finger.

Use of CAASPP Components

The LEA-4 CAASPP coordinator and an assistant superintendent described their attendance at the 2018 Fall CAASPP Institute as very useful. They both were impressed with the enhancements to Smarter Balanced components and were supportive of the LEA mandate for all schools to use IABs and ICAs. They also encouraged use of the DL. However, use of IABs and ICAs results was not extensive.

LEA-4 required all schools to administer NWEAs as benchmark exams. The mandated NWEA assessments were administered in the fall and spring for math, reading, and language at each school. Teachers and LEA leaders described varied purposes of the NWEAs including measuring growth, providing input to student classroom grades, and giving a sense of how well students were going to do on the CAASPP. The LEA assistant superintendent noted, "NWEA is complex. There are a lot of different reports. It's a firehose of complex data. If you use it correctly and have the time, you can do a lot of targeted work with students, but we don't have a reservoir of available time."

In addition to Smarter Balanced and NWEA assessments, LEA-4 also trained and encouraged PLCs to create assessments targeted to the essential standards using Illuminate, which was the district's vehicle for report cards. Teachers reported that some PLC meetings were used to create or select common assessments (e.g., from the Illuminate item bank, Go Math, Engage New York) or to evaluate different ways of teaching the same content by reviewing classroom assessment results. LEA-4-HS administered quarterly third party CCSS formative mathematics assessments (e.g.,

Lessoneer™, the Silicon Valley Mathematics Initiative), in addition to CAASPP components, to help identify best teaching practices. For ELA, teachers are constructing formative assessments in their PLC.

Summative Assessments

Access to and Use of Summative Data

LEA-4's approach to summative data is to hold off on data sharing and analysis until the CDE releases official results. Though the LEA shared embargoed data with school administrators in September, it directed them not to share those data yet with teachers. The LEA seemed uncertain about what was allowable prior to the official release. "There must be some reason that the state provides data to the district and then waits to release the official scores to the public. What this makes me think about is, for us, working on access to the data…I know we load the results in Aries and Illuminate. Even though it's in there, do teachers really know how to access them? That's something for us to work on." The Online Score Reporting System (ORS) was not used at the school level. Educators at the schools would have liked to have access to summative data sooner.

At LEA-4-ES, the principal conducted a half-day meeting with teachers to review and discuss 2018 summative results. Teachers were given a spreadsheet, prepared by the school's CAASPP coordinator, with scores of current and prior year students. Working in groups by grade level, teachers filled in blanks on graphic organizer posters with data (i.e., number of students tested, percentages of students at each achievement level and claim score), identified areas of strength/weakness, and noted steps in an action plan to improve scores. The posters were returned to the district, which presented them to the school board. Examples of action plan steps for grade three included:

- Math: (a) Use iPad apps/websites to reinforce skills in weak areas, (b) use manipulatives to reinforce concepts, (c) use strategic partners and groups, (d) use task cards and math talks.

- ELA: (a) Use practice tests from CAASPP website, (b) use Illuminate website to create standards-based tests, (c) teach the R.A.C.E.S. strategy to improve written responses, and (d) use Google classroom to improve written responses and typing skills.

At LEA-4-ES, grade level PLCs described several additional uses of summative data. For example, the grade four team compared their students' grade three summative scores with other measures (e.g., daily assessments) and used this information to group students for reteaching skills and spiral review. Spiral review is a process of revisiting topics repeatedly. Some teachers compared year-to-year results for individual students to identify areas of growth or lack of growth, and then reflected on their prior year's instruction to discover what worked.

At LEA-4-MS, the principal adhered to but expressed dissatisfaction with the district's guidance about how and when to use the summative assessment results. "We're told

the scores are not official and might change, so we shouldn't use them. I'm concerned about the data given to us so late. It should be official in August, if not sooner." The principal conducted a data analysis meeting in October with all teachers, including those for social studies and science. Teachers reviewed the following ELA and math results for grades six (incoming feeder students) through eight: bar charts that compared 2017 and 2018 year-to-year scale scores, achievement levels (also 2015 and 2016 for grades seven and eight), and claim scores to statewide averages; 2017 and 2018 rank ordered lists of middle schools by ELA and mathematics achievement level percentages; 2018 target score reports; and fall 2017 action plans for improving scores (from 2017–18).

Graphics of cohort data (e.g., 2017 grade seven compared to 2018 grade eight ELA scores) were not part of the presentation, though a few teachers mentioned they compared the performance of a cohort during an earlier data meeting. Teachers worked in grade level and subject area groups to develop 2018–19 action plans, with social studies teachers joining the ELA groups and science teachers joining the ELA and mathematics groups. This was the first year that teachers were given target scores, but teachers did not describe attending to them. The principal remarked, "There needs to be an analysis piece inside of this to tell us what to target. It's overwhelming to our teachers. If there was a diagnostic tool, that would be neat. All our [target] scores [areas where performance indicates standard met] are in the red [below standard met]." The LEA-4-MS mathematics PLC used Illuminate to create group reports of CAASPP scores for each teacher.

At LEA-4-MS, the 2018 summative scores were used in combination with other assessment results (NWEA) to identify and invite students to attend one-hour before-school tutoring and four-hour Saturday tutoring sessions. Teachers stated the aim of tutoring was to improve skills of 'the 'bubble kids' [20 grade seven and 34 grade eight students] that were right below 'at standard'. This approach differed from the prior year, when student identification for tutoring was based on grade point average. Due to union restrictions, tutoring was done by teachers who volunteer their extra time. The school offered additional sessions that any student could attend.

At LEA-4-HS, the principal reviewed the embargoed data and compared his school's performance to other schools. With only one year of high school data available, he found this of limited use. However, he acknowledged the reports provide a snapshot for staff, the school site council, the ELAC, and parents. "The way the results are reported out to parents has been beneficial when they come in to talk to us about their students' learning. The reports are written in a very parent-friendly model." The principal used ORS and the student information system to create a spreadsheet of grade eleven 2017 and 2018 results and shared data with teachers at an October data meeting. Achievement level and claim scores were examined, but not scale scores or target scores. The principal made attempts to match grade eight and grade eleven results for individual students to help interpret scores. "For ELA, we continued the same program and the same PD [professional development]. We added in more depth in what we were doing with ELA, but we saw a decline in our ELA score. Since I can't see how last year's grade eleven students did as seniors, I don't know if I had a really bright group of 11th graders or if my current group of 11th graders just didn't do as well."

At LEA-4-HS, grade eleven ELA and mathematics teachers continued reviewing summative assessment results during a PLC meeting and used claim scores to help identify areas of weakness and strength. For ELA, teachers found little differentiation in performance between the four claims but decided to focus their goals for improvement on writing and reading. Teachers noted students often did not see how writing assignments in different classes relate to each other. Their goals for the year included finding support in the textbook for specific lessons on the writing process and reviewing last year's lessons on the writing process to see how to improve them. The mathematics PLC team used the data to inform common lesson planning and creation of common assessments. Results from common assessments were used to determine how to provide appropriate student-level enrichment or intervention.

Also at LEA-4-HS, counselors used grade eight ELA results from feeder schools to help make decisions about student placement into the grade nine and ten ELA support classes; however, NWEA results were used to place students into the grade nine math support class. CAASPP Early Assessment Program (EAP) results typically arrive after decisions about student placement into grade twelve "college transition" courses have been made; however, counselors may use EAP scores to adjust enrollment in these courses during the school year.

Student Preparedness

LEA-4 did not conduct student focus groups. The information on student preparedness is from the perspective of school staff.

LEA-4-ES's principal described the challenge of targeting instruction at the rigor of CAASPP when students were not on grade level, and he stated teachers are unsure of what appropriate scaffolding looks like. "As teachers plan their essential standards or common assessments, those conversations come up ('I don't think they're ready yet'), but we need to figure out how to get them [students] through grade-level content." Teachers used various methods to prepare students for the demands of the summative assessments. Some teachers used CCSS-aligned reading activities (e.g., Newsela, ThinkCentral) once a week to give students practice answering high-rigor questions. *Read 180* was also used for intervention.

One teacher described the value of the online CAASPP Practice Test to prepare third grade students for using the tools on the summative assessments. Also, one teacher stated "Last year, I printed out the practice test because students don't have access to computers. I stapled the exemplar answers to the back of the practice test. They take it home over spring break and do the practice test with their parents…Also, something I did last year that I might do this year is having students work on practice tests in groups (giving groups of students 3-4 questions and have them explain their answers). That's really helpful to them."

LEA-4-MS PLCs created common assessments (benchmarks) in ELA and mathematics for grades seven and eight. ELA teachers created tests in Illuminate that targeted standards within an instructional unit and used a common rubric available from Smarter Balanced. Math teachers used some tasks from Mathematics Assessment Resource

Services (MARS) to create items that used the academic language of CAASPP. Math teachers emphasized they intentionally included more tasks, items with multiple answers, and longer problems based on groups of items. The tests were administered using Classkick, which teachers described as an interactive white board similar to Google Classroom. Teachers accessed the assignment (assessment) any time to control what students saw.

Interim Assessments

General Interim Assessment Information

LEA-4 staff and school principals established the assessment calendar for 2018–19. The LEA requires elementary schools to give the ICA in ELA and mathematics (January), the middle school to give one IAB (December), and the high school to give two IABs (November and January/February). The LEA-4-HS principal valued the ICAs as predictors of summative performance but lobbied the district to mandate IABs to avoid the long ICA testing time. The mandated ICAs and IABs must be administered in a standardized manner, but PLCs may choose which IABs are given. Additional IABs may be administered, and the manner of administration may be standardized or nonstandardized. In selecting which mandated IABs to take, teachers considered which test aligned to the sequence of instruction, which test would collect more information on an area where summative scores were weak, and which test would measure essential standards for the grade level.

At the three LEA-4 study schools, some teachers were unaware of IABs and other teachers expressed a preference for administering IABs over ICAs. The latter teachers said IABs better inform the need for reteaching specific skills, due to the reduced test length and quick return of results; also, they thought lengthy ICAs took away instruction time without providing results quickly enough to make instructional adjustments. Teachers also mentioned that some content on the ICA, when given in the mandated month, would not yet have been covered in class. At LEA-4-HS, some teachers were unaware the IABs were not adaptive, and others thought there were multiple forms of each IAB to better support pre/post testing.

Table C15 presents the total number of IAB tests taken during the 2018-2019 school year by students in the three LEA-4 study schools and the number of IAB tests taken by students in all schools in the LEA. Counts of tests include those for students who took the same test more than one time. The table also indicates how many enrolled students in the LEA and each school are eligible to take the CAASPP Summative Assessments. And, the table indicates how many of the total IABs were in each domain (ELA or mathematics) and how many were given in a standardized manner versus a nonstandardized manner. The table also indicates how many students enrolled in the LEA are eligible to take the CAASPP Summative Assessments.

Table C15. Number of Smarter Balanced IABs Taken by LEA-4 Students

	CAASPP Eligible Students	Total IABs	ELA IABs	Math IABs	Standardized IABs	Non-Standardized IABs
LEA-4	2,656	7,078	2,618	4,460	5,326	1,752
LEA-4-ES	332	805	319	486	803	2
LEA-4-MS	774	1,585	380	1,205	1,362	223
LEA-4 HS	344	2,611	1,132	1,479	1,825	786

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-4 had 2,656 students eligible for the CAASPP summative assessments in 2018–19. LEA-4 gave 7,078 total IABs (count of tests given). Of these, 2,618 tests were for ELA and 4,460 tests were for math. Of the total IABs, 5,326 were given in a standardized manner and 1,752 in a nonstandardized manner (across ELA and math).*

Table C16 presents the count of testing opportunities (i.e., test sessions) there were for specific ELA and mathematics IABs, by grade level, across the three LEA-4 study schools during the 2018–19 school year. Frequency of administration of an IAB for some grades includes more than one school. IABs that require hand scoring are noted in the table with an asterisk (*). The greatest number of ELA IABs (18), mathematics IABs (29), and total IABs (47) were given in high school. The most frequently given ELA IABs were Research and Revision (six times each in high school). The most frequently given mathematics IAB was Interpreting Functions (10 times in high school). In the table, NA indicates the IAB is not available at that grade.

Table C16. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-4, by Domain and Grade

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	HS
ELA	Brief Writes*	0	1	3	0	0	0	1
ELA	Editing	0	1	3	1	0	0	1
ELA	Language and Vocabulary Use	0	1	1	0	0	0	1
ELA	Listen/Interpret	0	1	3	0	0	1	1
ELA	Read Informational Texts*	3	0	1	0	0	1	2
ELA	Read Literary Texts*	1	1	3	1	1	1	0
ELA	Research	0	0	1	0	1	0	6
ELA	Revision	0	0	0	0	0	0	6
ELA	Performance Task*	0	0	0	0	0	0	0
ELA	**SUBTOTAL all ELA IABs**	**4**	**5**	**15**	**2**	**2**	**3**	**18**

Table C16. (cont.)

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	HS
Math	Measurement and Data	0	0	0	NA	NA	NA	NA
Math	Number and Operations - Fractions	1	1	1	NA	NA	NA	NA
Math	Number and Operations in Base Ten	1	0	1	NA	NA	NA	NA
Math	Operations and Algebraic Thinking	3	0	1	NA	NA	NA	NA
Math	Geometry	0	0	3	0	1	1	NA
Math	Expressions and Equations	NA	NA	NA	1	1	NA	NA
Math	Ratios and Proportional Relationships	NA	NA	NA	3	1	NA	NA
Math	The Number System	NA	NA	NA	1	1	1	NA
Math	Statistics and Probability	NA	NA	NA	0	1	NA	3
Math	Expressions and Equations I	NA	NA	NA	NA	NA	1	NA
Math	Expressions and Equations II	NA	NA	NA	NA	NA	1	NA
Math	Functions	NA	NA	NA	NA	NA	1	NA
Math	Algebra and Functions I	NA	NA	NA	NA	NA	NA	6
Math	Algebra and Functions II	NA	NA	NA	NA	NA	NA	6
Math	Geometry and Right Triangle Trigonometry	NA	NA	NA	NA	NA	NA	1
Math	Interpreting Functions	NA	NA	NA	NA	NA	NA	10
Math	Seeing Structure in Expressions/Polynomial Expressions	NA	NA	NA	NA	NA	NA	1
Math	Geometry Congruence	NA	NA	NA	NA	NA	NA	1
Math	Geometry Measurement and Modeling	NA	NA	NA	NA	NA	NA	1
Math	Performance Task*	0	0	6	0	1	1	0
Math	**SUBTOTAL all Math IABs**	**5**	**1**	**12**	**5**	**6**	**6**	**29**
BOTH	**TOTAL IABs, ELA and Math**	**9**	**6**	**27**	**7**	**8**	**9**	**47**

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at each grade across the LEA-4 schools in the study. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-4 there no testing opportunities for Brief Writes at grades 3, 6, 7, 0r 8, one each at grades 4 and high school, and three opportunities at grade 5.*

Table C17 presents the number and types of ICAs given by the three case study schools in LEA-4 over the duration of the 2018–19 school year, in comparison to the ICAs given by all schools in the LEA. Two of the study schools did not administer any ICAs; all ICAs were administered in standardized manner.

The need for extra time for hand scoring was cited as a reason for not selecting certain IABs. At LEA-4-MS, however, comments about why IABs with hand scoring were avoided indicated a misunderstanding. The principal stated, "We were told that if we gave an interim assessment, but did not hand score the students' test, the students' summative score would be negatively impacted. My assistant principal and mathematics department chair were told by someone from CAASPP that this was the case." This may be the result of confusion with the fact that results from IABs with hand scored items are not available until the hand scoring is completed. At all schools, some teachers were unaware of the option to administer interims in a nonstandardized manner. A few had learned about this option last year and were planning to do some group administrations this year. "I am interested in the idea of using the interims as a teaching tool… If students have a problem with directions and process, for me, it seems it would be better for us to take some time and go through the steps."

Table C17. Summary of Smarter Balanced ICAs Administered by LEA-4

	CAASPP Eligible Students	Total ICAs	ELA ICAs	Math ICAs	Standardized ICAs	Non-Standardized ICAs
LEA-4	2,656	1,140	548	592	1,048	92
LEA-4-ES	332	604	314	290	604	-
LEA-4-MS	774	-	-	-	-	-
LEA-4-HS	344	-	-	-	-	-

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-4 had 2,656 students eligible for the CAASPP summative assessments in 2018–19. LEA-4 gave 1,140 total ICAs (count of tests given). Of these, 548 tests were for ELA and 592 tests were for math. Of the total ICAs, 1,048 were given in a standardized manner and 92 in a nonstandardized manner (across ELA and math).*

Purpose of Interim Assessment Use

LEA-4 used ICAs and IABs last year primarily to prepare students and teachers for the summative assessment format, content, rigor, and interface. The goal of LEA-4 and administrators this school year was to encourage more teachers to use results to influence instruction.

At LEA-4-ES, the aim of ICAs was to help identify which essential standards students had not mastered, so instruction in these could be delivered before the Smarter Balanced summative assessments in ELA and mathematics are administered. ICAs

were scheduled for January 2019, so discussions with teachers were based on the use of ICA results during the prior year. During 2017–18, the CAASPP coordinator provided teachers with grade-level results by item; results were not teacher specific. Opinions about the usefulness of the 2017–18 ICA results varied. The principal thought the hand-scoring process was itself informative to teachers. They mentioned having a better understanding of how to "gear my instruction and teach what a constructed-response needs to look like in the context of my curriculum" and "where we need to [do] better (math performance task, multi-step problems)." However, teachers also thought the delay in getting results that confirmed what many of them already knew was not worth the lengthy ICA administration time.

At LEA-4-MS, the IABs were scheduled for December, so discussions were based on the IAB results from the prior year. Teachers voiced concerns about not having direct access to results – this suggests at least some teachers were not trained how to access data independently, which they should have had access to., The CAASPP coordinator provided results for the January 2017–18 IAB administrations in April. These were grade-level results by item; results were not teacher specific. Opinions about the usefulness of the 2017–18 IAB results varied. Some teachers discussed with students "what things concerned them, what things were foreign to them, what things did they get, what things did they need help with" but others mentioned that, because the tests aren't allowed to impact student grades, there's less buy-in from students, who don't take it very seriously.

At LEA-4-HS, some ELA teachers gave the IABs only to comply with district policy and were not planning to use the results. However, the mathematics PLC planned to analyze results from a recent IAB to drive instruction and improve from past scores. Math teachers also planned to print and share IAB reports with students to show their areas of strength and weakness.

Interim Assessment Reporting System

The LEA CAASPP coordinator stated teachers were given access to the Interim Assessment Reporting System (IA Reporting System). LEA and school staff were not knowledgeable about and had not used many of the enhancements made to the system since it was launched in spring 2015. LEA staff reported having difficulties logging into the system but were not specific about what caused the issues.

At LEA-4-ES, the principal expressed a need for more district support to obtain detailed test results ("We need to dial down reports, like students need help with these particular skills"). The school CAASPP coordinator did not think teachers had access to the IA Reporting System, which he used to prepare question-by-question results for teachers by grade level.

At LEA-4-MS, only the CAASPP coordinator accessed IAB results in the IA Reporting System in 2017–18. He prepared and distributed to teachers a limited set of ELA and mathematics results for each IAB administered. Results were presented as a table (number of students tested, average scale score, and percentage below, near, above

standard) and aggregated by grade level. Interims had not yet been administered in 2018–19.

At LEA-4-HS, the principal and CAASPP coordinator hadn't used the IA Reporting System since it was launched and did not know it now offers item-level information. Teachers had access to the system to retrieve their own IAB scores but were not provided training or guidance. The principal did not intend to train teachers on use of the system in 2018–19 due to lack of time.

Digital Library

The LEA-4 CAASPP coordinator and principals reported that DL passwords were issued to all teachers; however, very few teachers reported accessing the DL or using any of its resources. Some teachers didn't know what it was. At LEA-4-ES, teachers started to explore the DL, but noted they had access to so many resources, they haven't found it particularly useful ("You click around and then it's forgotten about"). The principal of LEA-4-MS planned to "infuse DL lessons" into the intervention instruction during Saturday school. The intent was to have teachers search the DL for lessons related to standards for which students' scores on the NWEA assessments (given twice a year) were low. Teachers had just participated in DL training at LEA-4-MS, during which they had some difficulties logging into the system and hadn't yet had time to "dig into it." A teacher experienced with the DL last year remarked on differences in the DL filtering options this year. LEA-4-HS's principal explained an "equivalent" resource (PLC Learning Solution Trees) was purchased, which met the need for formative assessments and CCSS-aligned lesson plans.

LEA-5 Findings

In LEA-5 one elementary school and one middle school participated in the study. Table C18 summarizes the qualitative data available for analysis. As shown, LEA-5 was a contributor of information to this study throughout the duration of the project. The LEA POC participated in the site visit interview and one initial monthly polling but did not contribute to other data collection activities.

Table C18. Summary of Data Sources for LEA-5

Data Source	
Site Visit Focus Group	ES – 4 teachers in one focus group: grade three (1), grade four (2), and grade five (1) MS – 3 teachers across three interviews: ELA (2), and math (1)
Site Visit Interview	ES – School POC (CAASPP Coordinator) and Principal, 1 interview MS – School POC (CAASPP Coordinator) and Assistant Principal LEA – LEA POC (CAASPP Coordinator)
Monthly Polling	ES – December (POC), January (12 teachers), February (POC, 2 teachers), April (3 teachers) MS – December (POC), January (13 teachers) February (POC, 2 teachers), LEA – December (POC)
End-of-Year Virtual Focus Groups	ES – POC MS - POC
Student Focus Groups	ES – 8 students MS – 6 students
Documentation	• Professional Development Agendas • CAASPP Interim Assessment presentation slides • Standards, Claims, and Targets documentation • School Interim Assessment schedules • Interim Assessment Fractions Analysis template for analyzing math IAs • Summative Assessment Blueprints • School Plan • Sample Lesson Plan • Summative Assessment data summary

LEA and School Characteristics

LEA-5 is a large district in southern California, with nearly 800 schools. Table C19 presents a summary of demographic and achievement characteristics for students in the LEA and the two study schools. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest). As shown, the schools' student populations consisted almost entirely of students who were socioeconomically disadvantaged. Both schools had a modest percentage of students with disabilities. The LEA-5-ES population has a large proportion of English learners. Both schools had a relatively small proportion of students that met or exceeded ELA or mathematics standards, although LEA-5-MS achievement rates had grown over the past couple years.

In different data gathering events, LEA-5-ES teachers described students as not possessing the reading, math, or computer skills necessary to succeed on the summative assessments. Teachers worried that students could not comprehend the test itself, even with the accommodation of having the test read to them. Across both schools, teachers believed not all students had computer experience or had access to the internet at home.

LEA-5-MS draws students from across the district because of its designation as a magnet school. This had increased integration across communities. While there was an application process, students within the school's boundaries could not be refused, and acceptance into the school was not based on meeting any specific criteria.

LEA-5-MS is currently a STEAM (Science, Technology, Engineering, Arts, and Math) magnet school, having transitioned to this educational approach to learning two years prior to the study. Simultaneous to the transition to a full STEAM magnet school, about half of the teaching staff were new hires. Many former LEA-5-MS teachers preferred to teach in a traditional school and transferred out. Many administrative staff are also new within the past two years.

Table C19. Demographic Characteristics of LEA-5 and Its Participating Schools

Variables	LEA-5	LEA-5-ES	LEA-5-MS
Enrollment	621,414	643	1,459
% Socioeconomically Disadvantaged	81%	95%	96%
% Students with Disabilities	14%	12%	13%
% English Learners	23%	44%	14%
% Reclassified Fluent English Proficient	27%	18%	44%
% Met or Exceeded ELA State Standards	42%	33%	40%
% Met or Exceeded Math State Standards	32%	22%	22%

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA overall and for each of the LEA's schools in the study. The second column (from top to bottom) shows in LEA-5 there was a total enrollment of 621,414 in 2017–18. Of these students, 81% were socioeconomically disadvantaged, 14% were students with disabilities, 23% were English learners, and 27% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 42% of students met or exceeded ELA state standards and 32% of students met or exceeded math state standards.*

Professional Learning Communities and Professional Development

LEA-5-ES conducted weekly PLCs by grade level for the purpose of planning and sharing strategies. LEA-5-MS organized its PLCs by content areas within each grade. PLCs in both schools met to discuss IAB scores and determine how to adjust their lesson plans based on the data.

LEA-5 provided numerous trainings throughout the school year about CAASPP. Each year there were three mandatory trainings about summative and interim assessments that administrators and CAASPP coordinators attended. Additional voluntary professional development opportunities have been available from the LEA. For example, the LEA-5-ES leadership team participated in a LEA-sponsored training about the Digital Library. Some LEA-5-MS teachers participated in district trainings on unpacking the NGSS standards. The participants in these district trainings receive presentation slides to then share with their school's teachers.

Administrators in LEA-5-ES described the challenge of organizing teacher professional development to address the various district initiatives. Administrators received trainings on a variety of CAASPP resources but struggled to find time to share what they learned with teachers. Quick chats or emails listing resources and how to access were the primary means of passing along the information, with some resources receiving greater attention in school training sessions for teachers (some of which were voluntary). Administrators and teachers suggested training on the CAASPP resources would ideally be provided throughout the year to increase depth of understanding; however, they said there was no time to feasibly accomplish that.

Initiatives and Resources

Teachers at LEA-5-ES and LEA-5-MS described their curriculum resources as insufficient to prepare students for the summative assessments. Some teachers suggested curriculum materials covered only superficially the necessary skills, while others suggested the curriculum materials were not aligned with the assessment. As a result, they had purchased other materials (e.g., Achieve 3000, StudySync, Newsela) or sought out resources on the internet (e.g., North Carolina Office of Education, Engage New York) that they believed better mirror the types and rigor of questions students experience on the summative assessment.

Technology

LEA-5-ES recently acquired Chromebooks for every two classrooms to share, as opposed to a school-wide shared computer lab. This increased technology use in the classroom; however, difficulties remained in sharing across classrooms. LEA-5-MS described a shift in technology. Resources for other technology initiatives were reduced and technology scheduling changed to make time for the interim assessments. They increased the number of accessible Chromebooks, but scheduling and sharing the Chromebooks was a challenge. iPads were also used for interim assessments. There were sometimes technical difficulties when using them, though one teacher noted the issues were less and less frequent over time. Additionally, one teacher noted the iPad

screens were very small and required lots of scrolling to get through text on the IABs. Students complained of the difficulty of navigating the interface on the small screen.

Use of CAASPP Components

LEA-5 strongly emphasized the use of data to improve teaching and learning. Administrators and some teachers described using summative assessment results to evaluate their teaching practices and adjust their lessons. Student testing results are also shared with students and parents, with the middle school teachers setting goals for student achievement based on summative results. LEA-5 also mandated IAB use by grade. Teachers and administrators could request waivers from the district IAB schedule. In general, teachers and administrators indicated that mathematics IABs were administered more than ELA IABs.

LEA-5 used its data system to report summative results. This system loads student data by teacher, which teachers prefer. Additionally, the elementary school used Benchmark, DIBELS, and DAZE assessments to monitor student progress.

Summative Assessments

Access to and Use of Summative Data

LEA-5 used summative data at the district level to identify weaknesses in understanding concepts as indicated by results from a large number of students across the LEA. The LEA then chose two mathematics and two ELA IABs at each grade level and mandated their administration for the 2017–18 school year across its schools.

Summative assessment results were initially shared with teachers in LEA-5-ES as a group. These results provided insight into (a) growth in ELA and mathematics compared to the previous year's assessment results and (b) the school's "distance from three" (i.e., the difference between the school's average scale score and the cut score for achievement level three, "met standard"). The district provided teachers with more results, including results disaggregated by student and overall class and for both previous and incoming year students. At the class and student level, teachers examined the distance from "met standard" achievement ratings. Overall, the summative assessment data were at the claim-level, which administrators perceived as too broad to adjust specific instructional practices.

LEA-5-ES teachers used summative assessment data to a) examine their teaching and b) set goals with students. Teachers considered the instructional practices accounted for the assessment results. Additionally, teachers conducted data chats with students to present charts illustrating students' distance from three (i.e., from "'met standard''); this information also was shared with parents. Teachers requested the summative test's Lexile level be reported so they could assess whether students possessed the reading capacity to understand the test.

LEA-5-MS followed a similar pattern in how summative assessment data were shared and used. At the beginning of the school year, the whole school met to review an overall report on grade-level ELA and mathematics year-to-year growth and "distance from three" results. Administrators shared a goal for the coming year's growth on assessment results, typically a 5% increase in scores. Grade-level content PLCs would then meet to examine their grade results, examining both the claim- and target-level results. Teachers indicated they would also examine individual student results, both previous and incoming students, and set goals for students to share with them and their parents. Some teachers felt the level of assessment result detail provided was sufficient to review lesson plans and presentations and adjust them to achieve better outcomes. Others requested results at the CCSS standard level so they could better understand what concept needed remediation, though this indicates lack of knowledge of the summative assessment test design and purpose. The ability of the summative assessment to differentiate skill level effectively was questioned by teachers at LEA-5-MS who found great variability of student capacity within a summative assessment achievement level.

Student Preparedness

Administrators and students reported comfort and confidence among the LEA-5-ES students during their summative assessments. Administrators perceived students felt relaxed taking the summative assessment on the computer. Students suggested their coursework and IAB experiences prepared them well.

LEA-5-MS students and administrators reported confidence in taking the summative assessment. A student focus group suggested a greater level of comfort with the ELA section than the mathematics section, as some students experienced technological problems when using the mathematics IABs in advance of the summative assessment. An administrator indicated the ELA IABs were emphasized throughout the year before the summative assessments.

Interim Assessments

General Interim Assessment Information

LEA-5 mandated a minimum number of IABs to be administered, two IABs in each area, each semester. LEA-5 set a schedule for each grade of when and which ELA and mathematics IABs should be conducted. Recommended IABs were selected based on summative assessment scores – particularly those addressing areas of weakness based on district-wide summative data. Schools could submit waivers to the district if they did not want to use the recommended IAB. Teachers were free to administer additional IABs. In LEA-5-ES, some PLCs chose to do many more than the minimum two, while other PLCs did not choose so many.

Teachers and administrators expressed some dissatisfaction with LEA-5's mandate. Schools must either submit a waiver, or administer the two IABs selected by the LEA. Teachers found the timing of the selected IABs did not match their curriculum schedule – this required them either to adapt their planning or conduct IABs that weren't relevant

to the unit at the scheduled IAB time. A general challenge was that IAB access was provided only after the CAASPP coordinator received training from the district and shared that training with the schoolteachers – this limited the time frame permitting IAB use.

Table C20 presents the total number of IABs taken by students in the two LEA-5 study schools during the 2018–19 school year, and the number of IAB tests taken by students in all schools in the LEA. Counts of tests include those for students who took the same test more than one time. The table also indicates how many enrolled students in the LEA and each school are eligible to take the CAASPP Summative Assessments. And, the table indicates how many of the total IABs were in each domain (ELA or mathematics) and how many were given in a standardized manner versus a nonstandardized manner. No ICAs were administered in the study schools.

Table C20. Number of Smarter Balanced IABs Taken by LEA-5 Students

	CAASPP Eligible Students	Total IABs	ELA IABs	Math IABs	Standardized IABs	Non-Standardized IABs
LEA-5	262,099	1,411,434	613,489	797,945	680,095	731,339
LEA-5-ES	315	2,460	843	1,617	1,560	900
LEA-5-MS	1,422	6,686	2,583	4,103	4,419	2,267

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-5 had 262,099 students eligible for the CAASPP summative assessments in 2018–19. LEA-5 gave 1,411,434 total IABs (count of tests given). Of these, 613,489 tests were for ELA and 797,945 tests were for math. Of the total IABs, 680,095 were given in a standardized manner and 731,339 in a nonstandardized manner (across ELA and math).*

Table C21 presents the count of testing opportunities (i.e., test sessions) there were for specific ELA and mathematics IABs, by grade level, by the two LEA-5 study schools during the 2018–19 school year. Frequency of administration of an IAB for some grades includes more than one school. IABs that require hand scoring are noted in the table with an asterisk (*). Grade six gave the greatest number of ELA IABs (17), grade four gave the greatest number of mathematics IABs (28), and grade four gave the greatest overall number of IABs (43). The most frequently given ELA IAB was Read Informational Texts (10 times in grade eight). The most frequently given mathematics IAB was Expressions and Equations I (15 times in high school). In the table, NA indicates the IAB is not available at that grade.

Table C21. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-5, by Domain and Grade

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
ELA	Brief Writes*	3	1	1	1	0	0
ELA	Editing	1	1	0	3	1	0
ELA	Language and Vocabulary Use	1	1	0	3	0	0
ELA	Listen/Interpret	3	1	0	1	0	0
ELA	Read Informational Texts*	3	6	1	3	3	10
ELA	Read Literary Texts*	3	3	3	1	1	1
ELA	Research	1	0	0	1	0	1
ELA	Revision	1	1	0	1	1	0
ELA	Performance Task*	0	1	0	3	0	3
ELA	**SUBTOTAL all ELA IABs**	**16**	**15**	**5**	**17**	**6**	**15**
Math	Measurement and Data	3	3	0	NA	NA	NA
Math	Number and Operations - Fractions	3	6	1	NA	NA	NA
Math	Number and Operations in Base Ten	6	10	6	NA	NA	NA
Math	Operations and Algebraic Thinking	10	3	0	NA	NA	NA
Math	Geometry	3	3	0	0	0	1
Math	Expressions and Equations	NA	NA	NA	6	10	NA
Math	Ratios and Proportional Relationships	NA	NA	NA	1	3	NA
Math	The Number System	NA	NA	NA	1	3	6
Math	Expressions and Equations I	NA	NA	NA	NA	NA	15
Math	Expressions and Equations II	NA	NA	NA	NA	NA	1
Math	Functions	NA	NA	NA	NA	NA	3
Math	Performance Task*	1	3	0	3	3	1
Math	**SUBTOTAL all Math IABs**	**26**	**28**	**7**	**11**	**19**	**27**
BOTH	**TOTAL IABs, ELA and Math**	**42**	**43**	**12**	**28**	**25**	**42**

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at each grade across the LEA-5 schools in the study. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-5 there were 3 testing opportunities for Brief Writes at grade 4, one each at grades 4, 5, 6, and 7, and none at grades 7 and 8.*

Teachers reported administering IABs in both standardized and nonstandardized manners. Teachers noted ELA IABs can be more difficult to go through in a nonstandardized way as it's difficult to project the question and the text at the same time. IABs were administered in different settings (e.g., text to speech, supervised breaks, translated math problems) for students who had accommodation plans in their

IEP. Teachers did not prepare students for IABs beyond teaching the curriculum. LEA-5-ES teachers administered fewer ELA IABs than mathematics because of the hand scoring required for many of them. IAB hand scoring was found to be very time consuming. LEA-5-MS teachers sometimes circumvented the frustration with time intensive hand scoring by only scoring a high-, medium-, and low-performing student to get a sense for how general groups in their classes performed. Teachers suggested the hand scoring rubrics were too broad to be uniformly useful for each question. Additionally, hand scoring training was not thought to be adequate. Teachers noted that hand scoring must be completed before accessing any other IAB results.

Purpose of Interim Assessment Use

IABs were commonly used, particularly in the elementary school, as pre/post assessments of students' learning. This allowed teachers to monitor student growth and share those results with students. The pre-test results were used by teachers to determine how they would approach teaching the material, as well as be aware of which students showed a higher risk of falling behind. Some PLCs mapped out what each item in the IABs covered and determined whether the curriculum sufficiently addressed the material or whether outside resources were needed. Post-test results from the IABs were used to identify gaps in student learning and reteach the necessary skills. The middle school conducted a "Smarter Balanced boot camp" where they used the results of all the previous IABs to develop review lessons on specific skills. One teacher created her own post-test, modifying the IAB questions slightly, rather than use the same IAB test for the pre- and post-test.

The IAB assessments were also used to familiarize students and teachers with the summative assessment format. For example, teachers described analyzing the vocabulary, question and answer format, and tech skills involved in the IAB to familiarize students with these elements. Similarly, others used the IAB items as a model to develop items of similar types and language. Teachers tried different accommodations with students to see what helped them most.

Teachers shared the results with the students. Students got to see how they did on each item. Student results were kept over time, either through individual folders with results from each assessment, or on a data wall in the classroom.

Teachers described showing IAB items after the assessment and reviewing different strategies and approaches that could be used to solve the problem. Other teachers used IAB items as warm up problems in their classes. IAB results were used to group students to provide instruction according to their specific needs.

In thinking about the changes to classroom assessment resulting from CAASPP implementation, one teacher posed the question that multiple teachers considered to some extent, "Are we preparing them for the SBAC? Versus are they learning the material?" She wondered whether the push toward CAASPP resources focused so much on familiarizing students with the summative test that it left little room to focus on students' grasp of the material.

Interim Assessment Reporting System

Teachers shared some difficulties using the IA Reporting system. One challenge was that the teacher needed to create the class group within the reporting system using the student SSID numbers to view the class results. Teachers who did not understand this process were confused because their students' results were mixed with other students. Because of the difficulty creating groups in the system, some teachers avoided using the IABs for individual assessment and instead administered them nonstandardized as a class learning activity. Teachers expressed wishes for changes to the reporting system. For example, they would like the ability to compare students' answers and to see students' testing history across IABs. Additionally, teachers would like to see IAB results related to claims, targets, and standards to relate them more closely to the curriculum. Some of these requests illustrate teachers were unaware of the full capacity of the IA Reporting System and may need additional training.

Digital Library

None of the teachers reported using resources in the DL. They suggested the DL was difficult to navigate to find useful resources and they did not have the time to explore. The LEA CAASPP coordinator indicated that DL use was part of training for school administrators, and LEA-5-ES school leaders reported receiving this training. LEA-5-MS administrators reported making DL an emphasis; however, neither school conducted specific training on DL use. Those who examined it found that they already had other resources that better met their needs. Others found problems with links within the system.

LEA-6 Findings

LEA-6's participation included one elementary school, one middle school, and one high school. Table C22 summarizes the qualitative data available for analysis. Though there were some missing data, LEA-6 and its schools were active participants throughout the study period.

Table C22. Summary of Data Sources for LEA-6

Data Source	Participants/Description
Site Visit Educator Focus Group	ES – 6 teachers in one focus group: Multiple grades (3), grade three (1), grade five (1), and Special Education (1) MS – 5 teachers across two focus groups: grade six (2), ELA (1), science (1), and history (1) HS – 2 teachers in one focus group: ELA (1) and mathematics/science (1)
Site Visit Leader Interview	ES – Principal, POC (teacher) MS – POC (Instructional Administrator) HS – POC (teacher)

Table C22. (cont.)

Data Source	Participants/Description
POC Monthly Polling	ES – February (3 teachers, 1 teacher/POC); April (2 teachers, 1 teacher/POC) MS – January (2 teachers); February (POC, 2 teachers) HS – January (1 teacher/POC), February (1 teacher/POC, 1 teacher), April (1 teacher/POC), May (1 teacher/POC) LEA – January (1 POC)
End-of-Year Virtual Focus Groups	ES – POC MS – POC HS – POC LEA – POC
Student Focus Groups	ES – 9 students MS – 6 students HS – 5 students
Documentation	• Fall 2018 IAB Schedule • 2019 Student IAB and CAASPP Goal Sheet • High School IAB Action Plan and Assessment Data • Instructional Leadership Team Agendas and IAB Decisions • CAASPP Planning Documentation • General LEA-wide IAB Usage Information

LEA-6 became a participant of the Impact Case Study later than most districts, so HumRRO's site visit did not occur until January 2019. Also, monthly polling did not begin until February 2019; the other districts started responding to monthly polls in December 2019. The timing of the in-person site visit may have impacted the type of information we received. For example, teachers had more time to administer interim assessments, and the 2018 summative results may have been more difficult for them to recall.

LEA and School Characteristics

LEA-6 is a medium-sized district in central California. The district includes six elementary schools, three middle schools (one of which is a charter school), and two high schools. The county also has four nontraditional schools. Table C23 presents a summary of demographic and achievement characteristics for LEA-6 school district and its three study schools. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest). As shown, participating schools had very few English learners. Just under half the student population was classified as socioeconomically disadvantaged, which was less than the state average (62%). The percentage of students with disabilities in the middle school (17%) was higher than the state average (11%). In terms of achievement levels, results were mixed. For example, LEA-6-ES exceeded the district in meeting or exceeding ELA and mathematics state standards while LEA-6-MS fell below the district and LEA-6-HS exceeded the district's ELA standard but fell below in math. Mathematics achievement was low for LEA-6-MS.

Table C23. Demographic Characteristics of LEA-6 and Its Participating Schools

Variables	LEA-6	LEA-6-ES	LEA-6-MS	LEA-6-HS
Enrollment	3,926	287	336	475
% Socioeconomically Disadvantaged	47%	42%	46%	41%
% Students with Disabilities	13%	12%	17%	11%
% English Learners	3%	1%	1%	2%
% Reclassified Fluent English Proficient	4%	1%	7%	9%
% Met or Exceeded ELA State Standards	45%	48%	42%	49%
% Met or Exceeded Math State Standards	29%	35%	14%	26%

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA overall and for each of the LEA's schools in the study. The second column (from top to bottom) shows in LEA-6 there was a total enrollment of 3,926 in 2017–18. Of these students, 47% were socioeconomically disadvantaged, 13% were students with disabilities, 3% were English learners, and 4% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 45% of students met or exceeded ELA state standards and 29% of students met or exceeded math state standards.*

Professional Learning Communities and Professional Development

LEA-6 used an instructional leadership team to disseminate information about CAASPP. This team consisted of all principals and vice principals, at least one teacher leader per school, one teacher union representative, and a few special education teachers. The LEA's CAASPP coordinator worked directly with school coordinators and started a monthly IAB newsletter that included a feedback form where teachers could respond.

There was no PLC standardization across the participating schools. LEA-6-MS had a mathematics PLC because that school fell well below the state on the summative assessment scores. This PLC had identified a learning objectives pathway for each grade and found that students had significant knowledge gaps in mathematics. LEA-6-HS did not have a PLC; they had tried to start them but there was no built-in time for collaboration. The school had five PD days a year. ELA-6-ES also did not have official PLCs but held biweekly "mini PDs" to share information.

Initiatives and Resources

Overall, the LEA-6 study schools reported a high retention rate among teachers and an "older generation" of teachers, especially at LEA-6-ES. This, however, had not been an issue with respect to teachers keeping up with innovations and changes in instruction.

Because the study schools were small, their teachers tended to have more varied jobs compared to teachers in larger schools.

Technology

A common theme was the inadequacy of access to and familiarity with technology, especially with respect to the CAASPP testing platform. In LEA-6-ES, it was noted that "We think students know how to use computers, but a lot of them don't." Many agreed that LEA-6 schools and students had less access/exposure to technology than what was required for CAASPP testing. LEA-6 used Chromebooks and some teachers noted they had about "half a class worth of computers." As a result, LEA-6-MS purchased supplemental programs so younger students could have more access to technology which, in turn increased motivation among those students and more comfort among teachers in allowing students to use the technology. These programs include Freckle, Typing Agent, and Accelerated Reader. Chromebook-to-student ratios have narrowed over time (LEA-6-ES is about a 1:3 ratio currently) and schools have received a more adequate supply of mice and headphones. However, having access to Chromebooks at any given time remained problematic as scheduling shared use was difficult.

Use of CAASPP Components

Though some individuals used and were trained on CAASPP resources in LEA-6, HumRRO found a lack of consistent use across the participating schools. The LEA had chosen not to mandate training on how to use CAASPP resources, preferring to encourage use and allow schools and teachers to make decisions during the 2018–19 school year. This flexibility meant schools approached CAASPP resource use differently. The LEA continues to improve its understanding, training, and integration of CAASPP resources over previous years. For example, district leaders improved their process for rostering students into the IA Reporting System to ensure each teacher had access to interim assessment data for their students. They had also improved integration of CAASPP data into Illuminate, so schools are able to expand CAASPP use.

Summative Assessments

Access to and Use of Summative Data

At the district level, CAASPP summative data were used to prioritize which professional learning teachers should engage in. For example, LEA-6 mathematics scores are low, so district leadership discussed the idea of bringing in mathematics consultants to help teachers with the mathematics framework and curriculum. Doing this was still in early stages—the process was somewhat "disjointed." For example, the results of last year's summative mathematics assessments were identified as an area of need but the data could be further refined to determine with which aspects of mathematics students need the most help.

Getting the summative information to schools started at leadership meetings when principals got together and discussed both good and bad highlights. Principals' took

responsibility to distribute data to teachers. Meetings took place at schools and teachers discussed scores at their respective grade levels. LEA representatives indicated summative scores were not used to instruct teachers on how to do their jobs, but there were differing opinions about this at the school level.

In schools, getting summative data was not necessarily easy and interpreting results not always straightforward to teachers. For example, one teacher noted, “we don’t get a report from the previous class.” Another said, “you have to ‘scrounge’ for them.” Though previous years’ scores can be accessed through Illuminate, teachers were not always aware of this. A teacher from LEA-6-ES said, “I knew from another teacher that I could go into Illuminate to get my scores.” Otherwise, the district shared a breakdown of class scores at the start of the year—but only the scores of the previous year’s students, not those of current year students. At LEA-6-HS scores are widely disseminated to staff, but it was unknown if teachers used the data.

Teachers discussed accessibility, ease of use and usefulness of summative assessment scores for instructional purposes. Some felt the year’s summative scores were not necessarily a good indicator of the current student skill level. Some teachers used the Aeries Student Information System to retrieve basic scores to form classroom groupings, indicating scores aren’t provided to them directly. Teachers here had recieved adequate training on how to interpret these scores, however. Teachers also indicated they used previous year’s scores to help direct current year instruction, especially if there was a noticeable knowledge gap. For example, LEA-6-ES teachers noticed previous year’s students had low speaking/listening scores. After researching the issue, they determined the problem was a lack of typing skills. They decided to focus more instructional time on typing.

LEA-6-MS teachers felt CAASPP testing is more integrated with curriculum, and even though no one believed that they “teach toward the test,” summative results are instructive not only to teachers to find gaps, but to directly inform students. For example, one teacher identified students who were on the cusp of the next achievement level and met individually with them to set personal goals showing them how many more points they needed to move up to the next level. For students to see how close they are to the next achievement level worked as a powerful motivator for the summative assessment. Finally, understanding where their school, as well as where they, themselves, fell on the spectrum of state averages was instructive, leading to improvements through a mixture of increasing student motivation, ensuring an appropriate testing environment, and improving instructional practice.

Student Preparedness

Opinions about preparedness for the summative assessments were mixed. Generally, there seemed to be consensus across groups (administrators, students, and teachers) that (a) some students were prepared, (b) there was engagement with CAASPP testing, (c) the test was long, and (d) some testing elements were unfamiliar. For example, “students aren’t finishing the performance task in five minutes like they had in previous years. They’re really understanding it and spending the time they need.” Whether or not students felt they had enough time to take the test was dependent on grade. For

example, “[for] some grade levels there doesn’t seem to be enough time and there are other grade levels that are flying through it a lot faster than expected.” Some students seemed to be taking less time on the computer adaptive portion because they understood how to use the tools while others, particularly the higher grades, wished they had more practice on the Chromebooks. Students who had more practice on a computer were glad for it. Some teachers noted the platform increased engagement. “Talking to the students in my class, most of what I’m hearing is that they are excited about it. They were saying testing has been the best part of their school year.” Some teachers noted that IABs had helped students feel prepared. An LEA-6-ES student said, “Yes, I feel like I learned everything we needed. We felt like we practiced a lot. It seemed easy.” There were more positive comments from students about being prepared for ELA versus mathematics.

Not all student and teacher feedback, however, was positive. Some students needed more preparation technologically to feel comfortable on the Chromebook. An LEA-6-ES teacher said some teachers were frustrated because students were being tested on things they had been learning since kindergarten, rather than a focus on the current year. Students’ negative comments focused on test length, how the format and content of the test was different than what was experienced in the classroom, and specific components of computers use, such as drag-and-drop. Older students noted the mathematics assessment tested material they hadn’t studied in this school year. Finally, there was some confusion about whether the performance task was timed, and some rushed through it and wished they could have done better.

Interim Assessments

General Interim Assessment Information

Through participation in CAASPP Institute the LEA learned IABs could be used formatively compared to the previous years when they were to be used for benchmarking. At the district level, the CAASPP coordinator and other staff strongly encouraged increased IAB use, but were careful to avoid mandates and specificity in how schools implemented use; however, the LEA expressed the expectation that teachers would give a minimum of two IABs before the summative assessment. The LEA conveyed the idea that IABs were useful to grasp summative assessment expectations, student and teacher exposure to test item types, and to inform teachers of instructional gaps. Because schools were given independence by the LEA to determine how to administer IABs, implementation was dissimilar across schools. For example, LEA-6-MS had a highly structured strategy, planning which teachers would give which IAB and when.

Table C24 presents the total number of IAB tests taken by students in the three case study schools in LEA-6 during the 2018–19 school year, and the number of IAB tests taken by students in all schools in the LEA. Counts of tests include those for students who took the same test more than one time. The table also indicates how many enrolled students in the LEA and each school are eligible to take the CAASPP Summative Assessments. And, the table indicates how many of the total IABs were in each domain

(ELA or mathematics) and how many were given in a standardized manner versus a nonstandardized manner.

Table C24. Number of Smarter Balanced IABs Taken By LEA-6 Students

	CAASPP Eligible Students	Total IABs	ELA IABs	Math IABs	Standardized IABs	Non-Standardized IABs
LEA-6	1,976	8,051	3,950	4,101	5,499	2,552
LEA-6-ES	173	439	171	268	288	151
LEA-6-MS	334	2,513	1,346	1,167	1,977	536
LEA-6-HS	89	717	567	150	618	99

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-6 had 1,976 students eligible for the CAASPP summative assessments in 2018–19. LEA-6 gave 8,051 total IABs (count of tests given). Of these, 3,950 tests were for ELA and 4,101 tests were for math. Of the total IABs, 5,499 were given in a standardized manner and 2,552 in a nonstandardized manner (across ELA and math).*

Table C25 presents the count of testing opportunities (i.e., test sessions) there were for specific ELA and mathematics IABs, by grade level, across the three LEA-6 study schools during the 2018–19 school year. Frequency of administration of an IAB for some grades includes more than one school. IABs that require hand scoring are noted in the table with an asterisk (*). High school gave the greatest number of ELA IABs (21). Grade eight gave the greatest number of mathematics IABs (18) and the most total IABs (27). The most frequently given ELA IABs were Language and Vocabulary Use (6 times in grade seven) and Read Informational Texts (6 times in grade six). The most frequently given mathematics IAB was The Number System (6 times in grade six and eight). In the table, NA indicates the IAB is not available at that grade.

Table C25. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-6, by Domain and Grade

Domain	IAB Name	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	HS
ELA	Brief Writes*	1	0	0	0	0	0	2
ELA	Editing	0	0	0	0	1	0	3
ELA	Language and Vocabulary Use	3	0	0	3	6	0	3
ELA	Listen/Interpret	0	0	0	0	0	1	2
ELA	Read Informational Texts*	3	1	3	6	1	1	1
ELA	Read Literary Texts*	0	1	1	3	1	1	2
ELA	Research	0	0	0	0	3	3	3
ELA	Revision	0	0	0	0	0	0	1
ELA	Edit/Revise	NA	NA	NA	NA	0	3	3
ELA	Performance Task*	0	1	1	1	0	0	1
ELA	**SUBTOTAL all ELA IABs**	**7**	**3**	**5**	**13**	**12**	**9**	**21**

Table C25. (cont.)

Domain	IAB Name	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	HS
Math	Measurement and Data	1	0	0	NA	NA	NA	NA
Math	Number and Operations - Fractions	1	1	1	NA	NA	NA	NA
Math	Number and Operations in Base Ten	1	1	3	NA	NA	NA	NA
Math	Operations and Algebraic Thinking	1	1	1	NA	NA	NA	NA
Math	Geometry	1	0	0	0	0	3	NA
Math	Expressions and Equations	NA	NA	NA	0	1	0	NA
Math	Ratios and Proportional Relationships	NA	NA	NA	3	3	0	NA
Math	The Number System	NA	NA	NA	6	1	6	0
Math	Expressions and Equations I	NA	NA	NA	NA	NA	3	0
Math	Expressions and Equations II	NA	NA	NA	NA	NA	3	0
Math	Functions	NA	NA	NA	NA	NA	3	0
Math	Algebra and Functions I	NA	NA	NA	NA	NA	NA	1
Math	Performance Task*	3	1	1	1	0	0	1
Math	**SUBTOTAL all Math IABs**	**8**	**4**	**6**	**10**	**5**	**18**	**2**
BOTH	**TOTAL IABs, ELA and Math**	**15**	**7**	**11**	**23**	**17**	**27**	**23**

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at each grade across the LEA-6 schools in the study. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-6 there was one testing opportunity for Brief Writes at grade 3, none at grades 4–8, and two at high school.*

Table C26 presents the numbers and types of ICAs given by the three case study schools in LEA-6 over the duration of the 2018–19 school year, in comparison to the ICAs given by all schools in the LEA.

LEA-6-ES administered IABs in 2017–2018, mostly to provide students practice with computers. In 2018–19 the school was more focused on using the IAB to uncover knowledge gaps. Though the process was smoother, there were still issues with teachers not being trained on the system and not having access to proper data. For example, one teacher reported, “I still can’t see my math scores.” “It shows me students who aren’t in my class.” “It shows an eighth period I don’t have.” The focus of LEA-6-MS was for teachers to understand the IAB process, establish a proper testing environment, set testing expectations, develop an alternative schedule, and treat the testing as an important aspect of the school year. Similarly, at LEA-6-HS, the main goal was to simply encourage teachers to use IABs after a glitch the previous year that did not allow teachers access to their data, which discouraged IAB use. A focus at LEA-6-HS was to administer literacy instruction across disciplines and to have “other disciplines

administer IABs." For example, "English did research, history did informational texts, and science did listening and interpreting," in order for everyone to have "buy in."

Table C26. Summary of Smarter Balanced ICAs Administered by LEA-6

	CAASPP Eligible Students	Total ICAs	ELA ICAs	Math ICAs	Standardized ICAs	Non-Standardized ICAs
LEA-6	1,976	140	21	119	120	20
LEA-6-ES	173	55	21	34	47	8
LEA-6-MS	334	-	-	-	-	-
LEA-6-HS	89	85	-	85	73	12

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-6 had 1,976 students eligible for the CAASPP summative assessments in 2018–19. LEA-6 gave 140 total ICAs (count of tests given). Of these, 21 tests were for ELA and 119 tests were for math. Of the total ICAs, 120 were given in a standardized manner and 20 in a nonstandardized manner (across ELA and math).*

The use of hand scoring varied across teachers at LEA-6 schools. According to the LEA, "We recommended staying away from [IABs] that require hand scoring because that's an additional area of learning. We want to get them to the point where we could collaboratively look at rubrics and hand score, but we're not quite there yet." Despite this sentiment, some hand scoring was performed and teacher attitudes towards it were generally neutral. Some teachers were trained formally, others received informal training from other teachers, some were self-taught, and some wanted to be trained either formally or by a colleague. One elementary school teacher had been formally trained in hand scoring and chose to train his students to hand score their own items. He felt this provided a good opportunity for students to understand the scoring process, and what was required to receive full credit. He felt students satisfactorily picked up the process quickly and found it an effective activity. Overall, it seemed that teachers were interested in learning it to have another instructional tool at their disposal. For example, one teacher, "thinks the hand-scoring process is particularly useful for kids who do not spend enough time on an item." At LEA-6-HS they planned to use "hand-scoring training guides and exemplars for ELA teachers, especially for brief writes and performance tasks." As a result, most IAB's were given in the standardized manner as encouraged by administration. Some exceptions to this were using the IAB for reviewing difficult concepts. LEA-6-MS science students and LEA-6-HS ninth grade students were given nonstandardized IABs for instructional purposes, so students could discuss results in teams.

Finally, the LEA reported the need to do a better job communicating with new special education teachers about designated supports and accommodations. Students were often provided all designated supports and accommodations instead of focusing on what is appropriate. Because LEA-6's special education department was in constant

transition, the district realized the need for more training and wished there were more video tutorials available. More positively, LEA-6-MS reported that designated supports and accommodations were automatically transferred from IEP's to the CAASPP program, so students who took IABs that were on an IEP were able to use the required accommodations. LEA-6-ES also reported trying to narrow down the accommodations.

Purpose of Interim Assessment Use

LEA-6 used IABs to expose students and teachers to the CAASPP testing format and rigor, and to help schools and teachers identify instructional gaps, set grade-level expectations, monitor progress, improve student motivation, and review for the summative assessment. Because IABs were better at identifying needs of current students in real time, the LEA encouraged its schools through various methods (newsletters, direct emails, etc.) to use IABs without overly mandating how. Additionally, interim testing allowed teachers to learn skills required for success on the summative assessment. One LEA representative said, "We've had a lot of really rich discussions about the language of the test and how the language of test is not the same as that used in classroom, which has shifted conversations about students struggling with language of test and using that language in instruction. Also, teachers could see rigor of standards a bit more. With interim assessments, teachers and students could increase familiarity with the interface and test questions, which helped prepare them for the summative assessment."

At LEA-6-ES, a significant change took place from the previous year on how IAB use was mandated by the LEA. Previously, teachers were required to give four specific IABs three times per year to show changes at different points in the year, which turned out to be an unsuccessful tactic. One teacher said, "everybody did them, but no one was talking about them. This year, we were told we had to give two IABs (it didn't matter which ones) at some point over the year. Most [teachers] gave five or six throughout the year when the IAB lined up with their curriculum, so that was pretty successful." Another teacher noted, "A lot of our staff…chose to use the information from the IABs as a way of reviewing before the summative test." Teachers also noted that IABs revealed that students were not exposed to CAASPP's unique question style/format. So, IABs were not necessarily used to explore the misunderstanding of a concept, but rather, how to answer a type of test question that students were unfamiliar with. Overall, many agreed the IAB's biggest benefit was "letting students know what is actually expected of them."

In LEA-6-MS, every grade seven student was administered four IABs unless they were in a special education program. The IABs were administered in classes other than mathematics and ELA. Because IAB use was widespread at LEA-6-MS, some had used it for more than just summative assessment preparation. For example, one teacher found that some students with lower grades did better on the research IAB than students with higher grades. Based on students' IAB scores, this teacher found a few students to participate in AVID®, a program to prepare students for college readiness. With the flexibility the LEA has given to when IABs can be given, LEA-6-MS teachers aligned test administrations with the curriculum, resulting in changes to the rigor and content of instruction. Others used the IAB as a form of instruction, taking the IAB together as a class to discuss results. LEA-6-MS reported positive results using IABs

perhaps because of the collaborative effort among teachers, across grades, and through PLCs, particularly in math.

In LEA-6-HS, there was some discussion/concern about having to "teach to the test" because of CAASPP. This seemed to be particularly true of mathematics where teachers have needed to shift from their mathematics testing sequence of Geometry instruction to Algebra II. Additionally, there was little cross-year coordination using IABs. Some teachers noted a big gap between grades 8–11 for students not to have been exposed to IABs and many would like to see an increase in student exposure before testing in grade eleven. LEA-6-HS engaged in more coordinated IAB testing in previous years, but less in the 2018–19 year. LEA-6-HS teachers would have liked to see more IABs per subject, to retest with new content. Overall, LEA-6-HS indicated less enthusiasm for IAB uses but agreed that exposure to the test was the most important benefit of the IAB.

Interim Assessment Reporting System

The LEA-6 POC acknowledged that in 2017–18, they unintentionally had not assigned roles in the IA Reporting System appropriately to allow teachers to see results. Teachers were able to administer IABs and were frustrated when they could not see how their students performed. The problem was fixed for the 2018–19 year, and most teachers did not have issues with the system. However, some teachers reported they could not see certain scores and/or they received data for students not in their classes. Now that student and class reports are mostly available to teachers, many were accessing and using the information.

LEA-6-MS had a more systematized use of IAB reporting compared to the other two schools. For example, it used Illuminate chart results and this information was posted in the staff room. Results were discussed at staff meetings where there were open discussions about "what the data mean." Outcomes of those collaborative meetings resulted in laying the groundwork for instructional changes. Grade-level teams and individual teachers were trained to retrieve and analyze their own data. This gave teachers the ability to work with individual students to share scores, showing them how close they were to meeting the standards and setting goals for future testing. Some students reported that not all teachers shared IAB results with them.

Digital Library

There was considerable consensus that the DL has improved since it was first rolled out and that access, while still difficult, had gotten easier. The LEA acknowledged that using the DL had not been a focus for the district and the goal this year was to promote it in a stepwise manner to avoid resistance at the school/teacher level: "We're hoping to work with teachers to get them looking at their framework and fill in the gaps in the curriculum with Digital Library resources. We're not there yet in terms of doing that systematically, but that's where we see this going over a multi-year sequence." Additionally, the LEA recognized the technical challenges in using the DL (e.g., two logons, elements requiring additional downloads).

The DL, while useful, was not necessarily comprehensive. A teacher at LEA-6-ES attempted to use the DL as her curriculum but realized it had instructional gaps. Other teachers at this and other schools, while familiar with the DL, simply did not know how to use it and others had never used it. More were not using it than using it. Many teachers would like to learn how to link IABs to the DL, but they had received no training. A further obstacle was a lack of time to learn and use the library; other resources compete with it for attention. A middle-school teacher noted that it was difficult to know which source of material was best for which situation and that at LEA-6-MS, AVID was being promoted, as it had more resources and staff had been trained on its use. Another competitor that was frequently mentioned was Teachers Pay Teachers. One LEA-6-ES teacher said it had been a struggle getting colleagues to look at the DL instead of other sources. Finally, there was some indication by a high school teacher that the ELA content was not adequate at the high school level.

LEA-7 Findings

LEA-7's participation included three elementary schools and one K–8 school. HumRRO requested the school POC of the K–8 school focus on the three upper grades to represent middle school CAASPP use in the district. Therefore, we discuss this fourth school as LEA-7-MS. Table C27 summarizes the qualitative data available for analysis. LEA-7 joined the study in April 2019 and thus schools were available for limited data gathering events. However, the schools were responsive in participating in telephonic interviews, monthly polling, and end-of year-focus groups.

Table C27. Summary of Data Sources for LEA-7

Data Source	Participants/Description
Site Visit Interview	LEA - POC (CAASPP Coordinator) ES1 – School POC (Principal) ES2 – School POC (Assistant Principal) ES3 – School POC (Assistant Principal) MS – School POC (Assistant Principal)
School POC Monthly Polling*	ES1 – POC, 3 teachers ES2 – POC, 4 teachers ES3 – POC, 3 teachers MS – POC, 4 teachers LEA – POC
End-of-Year Virtual Focus Groups	ES1 – School POC (Principal) ES2 – School POC (Assistant Principal) ES3 – School POC (Assistant Principal) MS – School POC (Assistant Principal) LEA – POC
Documentation	• Data Analysis Tool • CAASPP Data Staff Meeting Agenda and Presentation Slides

*Monthly polling for LEA-7 conducted at one time point due to late participation.

LEA and School Characteristics

LEA-7 is a medium-sized school district in central California. The district includes 17 elementary schools, 5 middle schools, and 2 K–8 schools. Table C28 presents a summary of demographic and achievement characteristics for LEA-7 and its four study schools. Data were obtained from the 2017–18 School Accountability Report Card and the CDE's website (DataQuest). As shown, all the study schools had high majority disadvantaged populations. Additionally, the four schools all had high levels of ELs, with LEA-7-ES1 the highest (38%). In terms of academic achievement on the Smarter Balanced Summative Assessments, most students at the elementary schools are similar to the LEA as a whole in not meeting or exceeding standards in ELA or mathematics, whereas LEA-7-MS achieves moderately higher in both ELA and Math.

Table C28. Demographic Characteristics of LEA-7 and Its Participating Schools

Variables	LEA-7	LEA-7-ES1	LEA-7-ES2	LEA-7-ES3	LEA-7-MS
Enrollment	22,777	692	915	827	957
% Socioeconomically Disadvantaged	85%	93%	90%	96%	78%
% Students with Disabilities	12%	11%	9%	5%	10%
% English Learners	23%	38%	28%	35%	32%
% Reclassified Fluent English Proficient	15%	7%	6%	7%	24%
% Met or Exceeded ELA State Standards	30%	21%	30%	24%	55%
% Met or Exceeded Math State Standards	20%	13%	21%	22%	39%

*Aside from Reclassified Fluent English Proficient, Data based on 2016–17 school year

Explanation of table contents: *For each variable in the first column, the next columns provide information for the LEA overall and for each of the LEA's schools in the study. The second column (from top to bottom) shows in LEA-7 there was a total enrollment of 22,777 in 2017–18. Of these students, 85% were socioeconomically disadvantaged, 12% were students with disabilities, 23% were English learners, and 15% were reclassified fluent English proficient. Results from the 2017–18 summative assessments indicated 30% of students met or exceeded ELA state standards and 20% of students met or exceeded math state standards.*

LEA-7-ES2 typically enrolls approximately 50 students who are homeless or in foster care. LEA-7-ES3's enrollment is relatively stable in number, but many students move into and out of the school. While LEA-7-ES1 has the lowest enrollment of all the schools, its upper grades average 33 students per class. The LEA-7-ES3 school POC noted many students at the school dealt with severe trauma and emotional issues.

LEA-7-MS was a K–8 dual immersion school. About an equal number of native English and Spanish speakers were in each grade. The instructional day was split such that half of the classes are taught in Spanish and half are taught in English.

LEA-7-ES1 had a high rate of turnover among experienced teachers during the past year. The school hired more than 10 new teachers; all grade three teachers were new. LEA-7-ES3 teachers were typically long-term employees.

Professional Learning Communities and Professional Development

In all the schools, teachers within each grade level functioned as a PLC. The PLCs in LEA-7-ES1 examined summative result data at the student level to identify which students scored low and which achieved "Met Standard" ratings. In LEA-7-ES2 and LEA-7-MS, PLCs played a big role in planning the use of IABs and evaluating the IAB results to guide instruction.

LEA-7 provided training for teachers at the beginning of the year about Smarter Balanced interim and summative assessments. LEA-7-ES2 teachers had not participated in formal, external CAASPP training conducted by the CDE and its partners, but the school's CAASPP coordinator had. No staff from LEA-7-ES1, LEA-7-ES3, or LEA-7-MS had participated in a formal, external CAASPP training. Most CAASPP-related training occurred through informal email from the CAASPP coordinator.

Initiatives and Resources

LEA-7 had such a large population of new students from other countries that they developed a newcomer program. This program gave students an opportunity for intensive English language learning with the intent of integrating the new students into regular school activities.

LEA-7-ES3 was a recipient of a school improvement grant that enabled numerous special initiatives. As part of this grant, additional personnel had been hired. The school had also hosted Saturday boot camps for teachers in grades three through five. These boot camps oriented teachers to the CAASPP online assessments, examining the tools and navigation as well as different question types. LEA-7-ES3 also conducted a parent information night to share different resources about the CAASPP System, such as the practice test website, information about different question types, and log-in information to access training at home.

LEA-7-ES1 engaged in a number of events to build students' confidence before summative assessments. Fourth grade teachers talked to third grade students about testing and what to expect, and younger grade students visited older grade students to cheer them on.

Technology

LEA-7 increased the number of devices to achieve one-to-one computer access for grades three through eight. They also provided typing practice devices and programs

for grade three. The LEA CAASPP coordinator noted one challenge to conducting summative assessments was that Chromebooks must be reset in order to enable text-to-speech accommodation on the Smarter Balanced assessment platform.

Use of CAASPP Components

LEA-7 emphasized the use of CAASPP assessment data. In each studied school, teachers collaborated with students to set goals for the student's achievement levels on the 2019 Smarter Balanced summative assessment. PLCs in LEA-7-MS used summative assessment results to evaluate what areas of the curriculum they should improve upon. Teachers administered at least one IAB in ELA and one in mathematics during 2018–19, as mandated by the LEA. IAB use was particularly focused on familiarizing students with the summative assessment format.

LEA-7 uploaded all CAASPP assessment results data into their Illuminate system. The district did this to avoid what they see as a cumbersome rostering process in the CAASPP reporting systems, ORS and the IA Reporting System. Housing the data in Illuminate also allowed for data from other assessments to be in the same system. Thus, multiple measures could be accessed in the same system for each student. LEA-7-ES3 created a school dashboard combining data from multiple sources (e.g., CAASPP tests, district benchmarks). LEA-7-ES2 administered additional assessments that come from their curriculum. These assessments were given three times each year and are 35-40 questions long.

Summative Assessments

Access to and Use of Summative Data

LEA-7-ES1 had a "Data Day" at the beginning of the year to examine the school's growth on the summative assessment from the prior year LEA-7-ES1 also used data to recognize individual students who showed growth between 2017 and 2018, and for those who performed well on their summative results in 2018. Teachers and students reviewed the previous year's summative results (2018) and set goals for the coming summative assessments (2019).

LEA-7-ES2 convened a big staff meeting before school started to examine summative assessment growth across subgroups in ELA and mathematics. A second meeting was held in November to look at the five-by-five placement reports. These reports were available on CDE's Dashboard website and combined data from the current and prior years to produce a color-coded grid. The grid had five Status levels (Very High, High, Medium, Low, Very Low) and five Change levels (Declined Significantly, Declined, Maintained, Increased, Increased Significantly). Data presentations also included comparisons with similar schools in the LEA. Similar to LEA-7-ES1, teachers and students set CAASPP goals in advance of summative testing based on the previous year's summative performance bands and current year IAB results.

LEA-7-ES3 primarily used summative assessment results to guide student goal setting. Teachers reviewed with students the previous year's summative assessment results (or

universal screening data for third graders, who had no prior year results) and set realistic goals for the coming year's summative assessment. These goals were referenced throughout the year. Teachers did not use the results for instructional purposes but intended to do so in the 2019–20 school year.

LEA-7-MS provided student-level scores to teachers at the beginning of the school year. Scores were provided both for students from the previous year (2017–18) and incoming year (2018–19). The teachers were provided additional data from other assessments and were tasked with determining how to help each student progress. Teachers set goals targeting growth for the student to aim for and communicated those goals to the student and parents/guardians. Summative results were also used by PLCs to identify areas in which students struggled the previous year so teachers could make plans to address the challenging areas.

The LEA CAASPP Coordinator reported that summative results were used by some schools to make class placement decisions. These schools attempted to balance the levels of achievement of students within a given class.

Student Preparedness

LEA-7-ES1 teachers felt the use of IABs familiarized students with the computerized assessment. Students seemed more comfortable and confident logging on, manipulating the mouse, and navigating the assessments. LEA-7-ES2 teachers felt they were able to effectively predict what the summative would look like and prepare students accordingly. During focus groups, students from LEA-7-ES3 told teachers they recognized in the summative assessment the item types they had previously practiced in class. LEA-7-MS students also seemed comfortable in the online testing environment, both in how to use various tools and recognize the types of items.

Interim Assessments

General Interim Assessment Information

The LEA mandated that schools administer one IAB in ELA and one in mathematics during the course of the year. The LEA encouraged schools to administer more IABs and showed data indicating that higher IAB use was associated with individual scale score growth. School principals were free to add their own expectations of teachers beyond the LEA's minimum mandate.

The assistant principal at LEA-7-ES1 established a schedule for each grade that indicated which IAB should be given in which week. The schedule included each IAB in both ELA and mathematics that did not require hand scoring. The assistant principal reported that teachers indicated they appreciated the structure and support.

Administrators at LEA-7-ES2 encouraged the use of IABs, particularly as a supplemental assessment at the end of each instructional unit. Grade-level PLCs and individual teachers could decide which IABs to take and when. The fifth grade teachers developed a system in the months preceding the summative assessment to give an ELA and mathematics IAB on Monday and Tuesday of each week, receive reports and

guiding questions for their PLCs on Wednesday based on the IAB data from the school CAASPP coordinator, and then conduct whole class instruction examining particular items from the IABs at the end of the week. Third and fourth grade teachers administered IABs in a less uniform and more infrequent manner. The CAASPP coordinator intended to provide a calendar for each grade next year to map out which IABs would be useful at the end of which units of instruction.

LEA-7-ES3 set an expectation that teachers would administer three IABs in each content area before the summative assessment. All teachers achieved that minimum, and some grades elected to give more IABs. Teachers who participated in Saturday Bootcamps, offered by the school, tended to give more IABs. Teachers and administrators also found great use of the CAASPP sample item website for daily practice items.

LEA-7-MS followed the LEA mandate to give at least one IAB in ELA and one in mathematics. All classes administered at least that many, with some grades giving many more. The third grade PLC gave more IABs to familiarize their students with the Smarter Balanced assessment interface because the students had no previous summative assessment experience.

Table C29 presents the total number of IAB tests taken by students in the four LEA-7 study schools during the 2018–19 school year, and the number of IAB tests taken by students in all schools in the LEA. Counts of tests include those for students who took the same test more than one time. The table also indicates how many enrolled students in the LEA and each school are eligible to take the CAASPP Summative Assessments. And, the table indicates how many of the total IABs were in each domain (ELA or mathematics) and how many were given in a standardized manner versus a nonstandardized manner. The majority of IABs at LEA-7-ES2 and LEA-7-ES3 were given in nonstandardized manner.

Table C29. Number of Smarter Balanced IABs Taken by LEA-7 Students

LEA-7-ES1	355	1,396	666	730	841	555
LEA-7-ES3	393	725	65	660	63	662

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-7 had 11,979 students eligible for the CAASPP summative assessments in 2018–19. LEA-7 gave 28,687 total IABs (count of tests given). Of these, 10,529 tests were for ELA and 18,158 tests were for math. Of the total IABs, 10,684 were given in a standardized manner and 18,003 in a nonstandardized manner (across ELA and math).*

Table C30 presents the count of testing opportunities (i.e., test sessions) there were for specific ELA and mathematics IABs, by grade level, across the four LEA-7 study schools during the 2018–19 school year. Frequency of administration of an IAB for some grades includes more than one school. IABs that require hand scoring are noted in the table with an asterisk (*). Grade three gave the greatest number of ELA IABs (44) and mathematics IABs (43), thus administering the most total IABs (87). The most frequently given ELA IAB was Read Informational Texts (14 times in grade three). The most frequently given mathematics IAB was Number and Operations in Base Ten (17 times in grade four). In the table, NA indicates the IAB is not available at that grade.

Table C31 presents the number and types of ICAs given by the four LEA-7 study schools across the 2018–19 school year, in comparison to the ICAs given by all schools within the LEA. Only one of the study schools administered any ICAs, and it gave the 4 ICAs in nonstandardized manner.

Across all schools, hand scoring was reported to be a source of frustration, and teachers generally avoided giving IABs that required hand scoring. LEA-7-ES1 teachers avoided this type of IAB in 2018–19 but intend to use them in the coming year. One grade in LEA-7-ES2 unknowingly selected as its first IAB of the year one that required hand scoring; responses were not graded for a long time, and teachers were subsequently hesitant to use any more such IABs since they required so much work. LEA-7-ES3 teachers were frustrated they could not access IAB results until the hand scoring item was scored. There were also complaints that some topics were only assessed by IABs that required hand scoring (e.g., Read Informative Texts, Read Literature Texts).

Table C30. Count of Opportunities to Take Specific Smarter Balanced IABs in LEA-7, by Domain and Grade

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
ELA	Brief Writes*	1	2	4	0	0	0
ELA	Editing	4	4	2	0	0	0
ELA	Language and Vocabulary Use	4	6	5	0	0	0
ELA	Listen/Interpret	5	3	1	1	1	1
ELA	Read Informational Texts*	14	5	5	0	0	0
ELA	Read Literary Texts*	10	5	8	0	1	0
ELA	Research	2	3	1	0	0	0
ELA	Revision	2	3	3	0	0	0
ELA	Performance Task*	2	4	3	0	0	0
ELA	**SUBTOTAL all ELA IABs**	**44**	**35**	**32**	**1**	**2**	**1**

Table C30. (cont.)

Domain	IAB Name	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Math	Measurement and Data	8	1	2	NA	NA	NA
Math	Number and Operations - Fractions	9	7	6	NA	NA	NA
Math	Number and Operations in Base Ten	11	17	6	NA	NA	NA
Math	Operations and Algebraic Thinking	10	3	5	NA	NA	NA
Math	Geometry	4	2	7	0	1	1
Math	Expressions and Equations	NA	NA	NA	0	1	NA
Math	Ratios and Proportional Relationships	NA	NA	NA	1	1	NA
Math	The Number System	NA	NA	NA	3	1	0
Math	Performance Task*	1	3	2	0	0	0
Math	**SUBTOTAL all Mathematics IABs**	**43**	**33**	**28**	**4**	**4**	**1**
BOTH	**TOTAL IABs, ELA and Mathematics**	**87**	**68**	**60**	**5**	**6**	**2**

Explanation of table contents: *For each IAB named in the second column, the next columns show how many testing opportunities (i.e., test administration sessions) there were at each grade across the LEA-7 schools in the study. The number of students who participated in each testing opportunity varied and may have been a full classroom of students or a select group of students. Row 1 shows that in LEA-7 there was one testing opportunity for Brief Writes at grade 3, two at grade 4, four at grade 5, and none and grades 6 through 8.*

Table C31. Summary of Smarter Balanced ICAs Administered by LEA-7

	CAASPP Eligible Students	Total ICAs	ELA ICAs	Math ICAs	Standardized ICAs	Non-Standardized ICAs
LEA-7	11,979	8	5	3	2	6
LEA-7-ES1	355	4	4	-	-	4
LEA-7-ES2	369	-	-	-	-	-
LEA-7-ES3	393	-	-	-	-	-
LEA-7-MS	612	-	-	-	-	-

Explanation of table contents: *The first row shows data for the LEA overall, and the next rows show data for each of the LEA's schools in the study. Row 1 shows LEA-7 had 11,979 students eligible for the CAASPP summative assessments in 2018–19. LEA-7 gave 8 total ICAs (count of tests given). Of these, 5 tests were for ELA and 3 tests were for math. Of the total ICAs, 2 were given in a standardized manner and 6 in a nonstandardized manner (across ELA and math).*

Purpose of Interim Assessment Use

Teachers at all schools emphasized the importance of IABs in familiarizing students with the computer assessment environment. Teachers noted the need for students to learn how to log on, use accommodations, use test tools, and become familiar with the format of the assessment and its items. Teachers from the three elementary schools indicated they had used IABs to help students understand the types of items on the assessment. They also used Smarter Balanced sample items as practice questions. Teachers commented that examining the items also helped them become familiar with the skills they should target in their instruction. One teacher described it this way, "I think it sets a nice tone of high expectation and rigorous answers that are expected of students." Teachers felt increased familiarity with item types would lead to improved outcomes on subsequent assessments. Additionally, teachers felt taking the IAB helped students practice the behavioral expectations of testing.

LEA-7-ES2 and LEA-7-MS teachers used the results from the assessment to guide instructional decisions. For example, teachers would review the results and construct small groups to provide differentiated instruction according to student needs. Other teachers conducted item analyses to identify particularly difficult items and then developed whole group instruction around the concepts and strategies necessary for these items. Teachers also reported that IABs were used before the summative assessment to see which skills students still possessed from earlier instruction on the subject and which skills would need additional review.

LEA-7-ES3 teachers talked about using the results to assess the effectiveness of instruction. The school's CAASPP coordinator suggested IABs are a tool "to determine if instruction is effective" while teachers stated that IABs helped them see "where does instruction need to change."

Interim Assessment Reporting System

LEA-7-ES2 teachers and administrators reported that teachers were unable to obtain data for their classroom-level IAB results. The data in the IA Reporting system were by grade, not by teacher, which was contrary to some of the intended uses of administering the IABs. The school's CAASPP coordinator therefore obtained the teacher-specific data for the teachers. The CAASPP coordinator currently pulls each student's data and then creates class results on his own. Teachers have indicated they would like more access to data and to have a better view of item-by-item results, indicating at least some teachers were unaware that this information was available through the IA Reporting System.

LEA-7-ES3 teachers reported not accessing the IAB data using the IA Reporting System. It's unclear whether they used another mechanism, or they did not examine the data. Some teachers in LEA-7-MS were able to access results in the IA Reporting system while others were not.

The LEA CAASPP coordinator found it much easier to upload results into Illuminate. Classes are already set up in Illuminate, so teachers can more easily investigate their

class results. The LEA CAASPP coordinator was aware of and did not like the fact that this approach loses the links to the DL playlist.

Digital Library

None of the schools' teachers reported using DL resources. Teachers suggested they had numerous other resources (e.g., Benchmark, Eureka Math, NewsELA, Prodigy, Zearn) to help with lesson planning and formative assessment. Other teachers felt they had insufficient training to use the resource. The LEA administrator suggested that log-in difficulties associated with attaching roles when individuals self-register have slowed the adoption of DL. Administrators at LEA-7-ES1 indicated teachers reported being unable to log on to the DL or needing training on how to use the resource. Teachers at LEA-7-ES2 felt there were so many new curriculum resources to learn, they did not have time to learn the DL as well. LEA-7-ES3 teachers reported the system continued to have glitches, as in the previous year. One teacher from the district mistakenly thought the DL was a site for students to borrow e-books.

This page is intentionally blank.

Appendix D: Summaries of LEA-Specific Findings from Impact Case Study

The following sections present a summary of findings and conclusions for each LEA and its sample of schools. Each summary includes (a) an overview of the context of the LEA and its schools, (b) a summary of findings about usage of each of the three CAASPP components studied (summative and interim assessments and the DL), and (c) HumRRO's identification of several best practices in the use of CAASPP components. Among the LEAs studied, there were some variations in the grades at each school level. For example, some elementary schools had kindergarten through grade five, while others also had a grade six. For the study, HumRRO classified findings from schools consisting of middle grades between elementary (ES) and high school (HS) as middle schools (MS), though some were named junior high schools. All the schools classified as middle schools included grades seven and eight, with some variation in the lowest grade. Appendix C describes in much greater detail the specific findings from each of the seven LEAs, including contextual descriptions and distinctions between findings at the school level (elementary, middle, high school). Each LEA-specific section in the appendix is organized thematically by key topics of the research questions. The experiences described in this appendix may be useful to LEAs and schools across California interested in increasing their effective use of CAASPP components or in identifying ways to improve their implementation.

LEA-1 Summary of Findings and Best Practices

LEA-1 is a medium-sized district in central California with a majority socio-economically disadvantaged student population and a high percentage of English learners. Regarding academic achievement in 2018, nearly one-third of LEA-1 students met or exceeded the grade level standard in ELA and nearly one-fourth did so for mathematics. Two elementary schools and one middle school collaborated with HumRRO on this study and participated fully in data collection activities. LEA-1 teachers described having a strong PLC culture with regularly scheduled time for grade- and content-level meetings.

In addition to Smarter Balanced components, LEA-1 used many digital ELA, literature, and mathematics resources to enhance student learning related to the CCSS. LEA-1 used SchoolCity[4] to generate and administer formative and interim assessments. Teachers noted SchoolCity had a strong item bank and was user friendly for creating assessments, with test results indicating which students have mastered or need to be retaught specific skills. LEA-1 schools reported good access to technology, and no teachers expressed technology was a barrier to student success on the summative assessments.

Regarding CAASPP summative data, the LEA-1 CAASPP Coordinator downloaded data from ORS and uploaded data to SchoolCity for school staff to view. Each school

[4] SchoolCity was purchased by Illuminate in June 2018. In this report, we use the name used by the LEA.

had access to performance levels, scale scores, and claims data in SchoolCity and could view school-, grade-, and teacher-level results; however, some teachers found the results of limited value and others were not accessing results. LEA-1's CAASPP Coordinator indicated target reports, which were not available in SchoolCity, became useful in 2018–19 because areas that showed improvement could be used to help pinpoint what was working well. LEA and school staff, as well as students, generally indicated students were prepared for the 2019 summative ELA and mathematics assessments; however, some students reported feeling unprepared for the content and rigor of items.

LEA-1 did not mandate the use of IABs for the 2018–19 school year, due to prior negative responses to such mandates, but encouraged use by highlighting the value of IABs for assessing the CCSS and preparing for the summative assessments. Decisions about using IABs were generally made in PLCs. LEA-1's CAASPP coordinator, rather than site coordinators, created a roster for each teacher's class and facilitated teachers' access to IABs. Schools used standardized and nonstandardized manner of administration of IABs. Students in grade five, across the study schools, had the most opportunities (45) to take a variety of IABs (9 different IABs in ELA, 6 in mathematics). In addition to using IABs to familiarize students with all aspects of the platform (e.g., universal tools and the calculator), teachers cited examples of using IABs to inform instruction. For example, teachers selected IAB items for distributive practice, gave students experience applying higher-level thinking skills on IABs with constructed response items, and used IAB hand scoring experience to identify student weaknesses. PLCs also discussed IAB results to identify gaps in student knowledge and determine how to target instruction. Schools in LEA-1 typically viewed IAB results in the IA Reporting System.

LEA-1 and study schools' staff were aware of the Digital Library but rarely used it, citing sufficiency of other available resources.

HumRRO identified the following best practices in LEA-1 supporting effective use of CAASPP components to improve teaching and student learning:

- LEA and school leaders empowered teachers to work together to determine the best use of CAASPP components, and other available resources, to make instructional decisions.

- PLCs worked collaboratively to identify how CAASPP components would be used in their classrooms, and to interpret and use assessment data to identify student gaps and inform instruction.

- Schools offered students many opportunities to take computer-based assessments, including the IABs, SchoolCity assessments, and CAASPP training and practice tests. These experiences resulted in students feeling prepared for the technology format of the summative assessment.

LEA-2 Summary of Findings and Best Practices

LEA-2 is a small but growing district in southern California, with nearly half the student population socioeconomically disadvantaged and close to a fourth classified as ELs. In 2018, about 60 percent of students in LEA-2 met or exceeded the grade level standard in ELA and slightly less than half did so for mathematics. One elementary school participated fully in data collection activities, and a middle school and high school primarily collaborated with HumRRO for the site visit. LEA-2 offers formative and summative assessment and DL training via a train-the-trainer model. A small number of teachers received direct training from LEA-2, which they in turn shared with their school's staff. Teachers and administrators received training on CAASPP, but indicated some training sessions, consisting of more than 100 slides, were overwhelming.

In addition to Smarter Balanced components, LEA-2 schools adopted CCSS-aligned curriculum with embedded online practice assessments and performance tasks. Student access to computers varied. Elementary teachers used Chromebooks in the classroom, middle school teachers used the computer lab or supplemented classroom Chromebooks with those from a technology cart, and high school teachers shared Chromebook carts.

LEA-2 is working towards increased availability and integration of school-level CAASPP summative assessment data. Elementary teachers used data to examine curriculum by grade level (e.g., grade five needed more resources in writing, basic multiplication skills, and geometry); however, middle school teachers received only overall scale scores and achievement level data, and high school teachers could only access data for their current year students.

LEA-2 elementary schools began using IABs in 2017–18. This year, LEA-2 encouraged but did not mandate the use of IABs at all school levels. Across the study schools, students in grade seven had the most opportunities (46) to take a variety of IABs (9 in ELA and 5 in mathematics), including ones requiring hand scoring. No IABs were given in the high school; teachers cited reasons of insufficient time and extensive alternate resources. No ICAs were given in LEA-2. In addition to using IABs to teach students how to take Smarter Balanced assessments, third grade teachers administered IABs so students could practice typing, and teachers of other grades had students practice on areas of weakness or review rubrics to learn requirements for a thorough answer. Teachers were just beginning to use standardized IABs to measure student progress and inform instructional decisions; however, there were some challenges with hand scoring and accessing results.

Most teachers in LEA-2 did not use the DL. Those who accessed the DL found it difficult to navigate, though a few teachers gave examples of finding and using instructional resources.

HumRRO identified the following best practices supporting effective use of CAASPP components to improve teaching and student learning:

- LEA and school leaders recognized the importance of teacher and student familiarity with CAASPP and provided needed training, resources, and support so staff felt comfortable using interim assessments and students had practice with the tools and item types used on the summative assessment.

- Elementary school grade-level teams worked collaboratively to plan use of interim assessments to assess student progress, identify need to reteach, inform instruction, and familiarize students with summative assessment format, tools, and rigor.

- LEA and school leaders developed plans to increase use of CAASPP components gradually and strategically.

LEA-3 Summary of Findings and Best Practices

LEA-3 is a direct-funded charter high school in southern California, with a low percentage of students classified as ELs (7%) and a significant percentage classified as socioeconomically disadvantaged (60%). Students within LEA-3 boundaries are given first priority for enrollment, followed by students living within the local school district, followed by students outside those boundaries. While most students enter LEA-3 in the freshman year, a fair number of students also enter as sophomores and juniors. LEA-3 is a high achieving school; in 2018, over 83 percent of students met or exceeded the grade eleven standard in ELA and 73 percent did so for mathematics. LEA-3 participated fully in the site visit but had reduced participation in later data collection activities. Teachers were organized into subject- or course-level PLCs, which reviewed summative results, determined when interim assessments were given, and created end-of-unit assessments and associated practice tests. A grade level team (principal, counselor, academic advisor) interacted with course-level PLCs to monitor student progress.

In addition to Smarter Balanced components, LEA-3 used a variety of non-CAASPP resources, including embedded curriculum assessments (e.g., College Board's Springboard) to prepare students for the rigor and focus of the summative assessments and practice taking tests on a computer. In grades nine and ten, NWEA assessments were used to monitor student progress. The block schedule at LEA-3 resulted in some students taking the summative assessment the semester after taking the corresponding ELA or mathematics course. LEA-3 had a one-to-one ratio of students to computers, and no teachers expressed technology was a barrier to student success on the summative assessments.

Regarding CAASPP summative data, teachers received initial results—achievement level scores by subject and year-to-year grade level growth—in the fall from the CAASPP coordinator. Later in the year, PLCs met with the CAASPP coordinator, who provided teacher rosters with claims and target reports, to guide goal setting and instructional planning. Teachers acknowledged receiving training on the use of ORS but had little experience using the system independently.

LEA-3 strongly encouraged but did not mandate the use of IABs during the 2018–19 school year. PLCs were empowered to decide which IABs were given, and when. LEA-3

gave no ELA IABs, but grade eleven students had many opportunities (20) to take a variety of IABs in mathematics (9 different tests). IABs were administered nearly equally in standardized and nonstandardized manner. In addition to using IABs to familiarize students with all aspects of the platform (e.g., calculator), teachers cited examples of using IABs to monitor student progress (e.g., administer as end of unit test) or inform instruction (i.e., identify skills to focus on during daily warm ups). There were some challenges regarding hand scoring and accessing results. The Interim Assessment Reporting System was found to be difficult to navigate and teachers did not understand how some items were scored.

HumRRO identified the following best practices supporting effective use of CAASPP components to improve teaching and student learning:

- LEA-3 targeted specific software or curriculum to provide relevant online assignments and experiences to prepare students for rigor and format of the summative assessment.
- PLCs were empowered to decide among themselves which IABs to use for a given course (e.g., Algebra) and when to administer them.
- Teachers used IABs in a nonstandardized manner, identifying peculiarities in the interface and question types to make students comfortable with the summative assessment and help ensure results would reflect their knowledge and skills.
- Teachers used the IABs to examine how well the school's curriculum prepared students for the summative assessment.

LEA-4 Summary of Findings and Best Practices

LEA-4 is a small, rural district in an agricultural area of central California, with a vast majority of students classified as socioeconomically disadvantaged. Among its three study schools, the English learner population is highest at the elementary school (45%) and much lower at the middle and high schools (18% and 12%, respectively). Regarding academic achievement in 2018, most students in the study schools did not meet or exceed grade level standards in ELA or mathematics (18%–38% met or exceeded in ELA, 9%–31% did so for math). One elementary school, one middle school, and one high school collaborated fully with HumRRO on the site visit; however, LEA-4 terminated its participation in the study in January 2019, due to lack of time for school staff to participate in data collection activities. LEA-4 funded teams of principals and teachers in 2018 to attend PLC training sessions, and the district supports weekly early release time for structured PLC meetings. At the study schools, formal PLCs were in the early phase, with the challenge of many new teachers at the elementary and middle schools. A districtwide goal was for PLCs to identify grade-level "essential standards" in each subject area.

In addition to CAASPP components, the Northwest Education Association (NWEA) Measures of Academic Progress (MAP) test is administered twice a year, as mandated by LEA-4. Schools use multiple measures of student achievement to make decisions

such as placement in courses or interventions (e.g., tutoring or Saturday school). Teachers use third-party CCSS-aligned questions (e.g., from Illuminate) to create formative and diagnostic assessments, including common grade level assessments. Some teachers also used CAASPP components (e.g., Practice Test) to familiarize students with Smarter Balanced testing features.

Regarding CAASPP summative data, LEA-4 shared 2018 achievement level results (embargoed data) with school administrators in September and directed them not to share data with teachers until official results from CDE were released. Teachers reviewed 2018 results in October and November and used them to inform big-picture action plans by grade and subject area. Teachers expressed interest in receiving additional detail in their score reports and having access to results earlier.

For the 2018–19 school year, LEA-4 mandated the administration of and schedule for Smarter Balanced ICAs for elementary schools and IABs for middle and high school, with the goal of encouraging more teachers to use results to influence instruction. The mandated ICAs and IABs were given in standardized manner, but PLCs could choose which IABs were given and additional IABs could be administered. Students in high school had the greatest number of opportunities (47) to take a variety of IABs (7 different tests for ELA and 8 for mathematics).

Use of the Digital Library was extremely limited; most teachers said they were unfamiliar with this resource.

Though the study of LEA-4 concluded prematurely, HumRRO identified several best practices that resulted in the success of CAASPP component use:

- LEA and school leaders provided training and empowered teachers to collaborate in PLCs to plan curricula, review and analyze CAASPP Smarter Balanced and other assessment data, and use evidence to evaluate instructional practices.

- School leaders encouraged teachers' use of resources other than textbooks, including CAASPP Smarter Balanced components, to develop mastery of skills in the CCSS in ELA and mathematics.

- In selecting which mandated IABs to take, teachers considered alignment to the sequence of instruction, the need for collecting more information on areas where summative scores were weak, and measurement of standards identified as essential for the grade level.

- Some teachers worked individually and collaboratively to access and use IAB data to identify specific student strengths and weaknesses and make decisions about the next steps for instruction.

LEA-5 Summary of Findings and Best Practices

LEA-5 is a very large district in southern California. One elementary and one middle school collaborated with HumRRO on this study and participated fully in data collection activities. The LEA POC participated in the site visit interview and the initial monthly polling. The majority of students at both study schools (95–96%) were classified as socioeconomically disadvantaged. The elementary school had 44 percent ELs and the middle school 14 percent. In 2018, most students in the study schools did not meet or exceed grade level standards in ELA or mathematics (only 33%–40% met or exceeded in ELA, 22% met or exceeded in math). Teachers in the study schools met in PLCs by grade level or grade and content area.

In addition to Smarter Balanced components, the elementary school used a variety of assessments to monitor student progress (e.g., DIBELS & DAZE). To supplement curriculum and enhance student learning related to the CCSS, the study schools purchased other materials (e.g., Achieve 3000, StudySync, Newsela) and used internet resources (e.g., North Carolina Office of Education, Engage New York) that mirror the types and rigor of questions seen in the CAASPP summative assessments. Pairs of elementary school classrooms shared Chromebooks which, in turn, increased technology use in the classroom. Middle school teachers competed for use of a limited number of Chromebooks, and often used iPads to administer the IABs. One teacher expressed that the iPad and Chromebook screens are very small, which made navigating through text on tablets difficult for students.

Regarding CAASPP summative data, LEA-5 downloads student data for each teacher from ORS to the local information system, so teachers do not need to use ORS to access their students' results. Teachers use reports to guide student- and class-level goal setting. For example, teachers analyzed and gained insight from (a) growth in ELA and mathematics compared to the previous year's assessment results and (b) the distance of students' scale scores from the scale scores required to meet the state mathematics and ELA standards. LEA and school staff, as well as students, generally indicated students were prepared for the 2019 summative ELA and mathematics assessments.

LEA-5 mandated that schools administer two IABs each in each content area each semester in the 2018–19 school year. The LEA also specified which ELA and mathematics IABs should be given along with a schedule for their administration. LEA-5 selected the IABs based on district-wide summative assessment scores, targeting areas of known student weakness. Though schools could submit waiver requests to the district to use an alternate IAB, and teachers could administer additional IABs, there was some dissatisfaction at the school level with the mandated approach and how it was implemented. Across the study schools, students in grade six had the most opportunities (17) to take ELA IABs (9 different tests), while students in grade four had the most opportunities (28) to take mathematics IABs (6 different tests). Schools used standardized and nonstandardized manner of administration. In addition to using IABs to familiarize students with the summative format and how test content aligns with curricular materials, teachers used IABs to inform instruction. For example, elementary teachers gave IABs as pre/post assessments of learning, and the middle school used

the results to develop review lessons for "SBAC boot camp." Teachers found hand scoring time consuming and burdensome. Teachers found it challenging to use the IA Reporting System to obtain classroom level results.

None of the school teachers participating in the study reported using resources in the DL. They suggested the DL was difficult to navigate and that finding useful resources was time consuming.

HumRRO identified the following best practices supporting effective use of CAASPP components to improve teaching and student learning:

- Teachers used IAB pre-testing to investigate whether the curriculum for the coming unit covers the elements assessed in the IAB. They could then add to the curriculum as necessary.
- PLCs worked collaboratively to use IAB post-test results to develop review lessons targeted to improve the weakest skills.
- Teachers shared summative and IAB results with students to keep them abreast of their progress and point out areas where they needed improvement.
- LEA conducted mandatory and voluntary CAASPP training sessions in the use of Smarter Balanced components, including the DL, and provided presentation slides for sharing information at the school level.
- LEA strongly emphasized the use of data to improve teaching and learning and provided summative results rostered by teacher to support evidence-based decision making.
- Though LEA-5 mandated specific IABs schools had to administer, it offered flexibility through a waiver request process, allowing teachers to ensure the IABs aligned with their classroom instruction.

LEA-6 Summary of Findings and Best Practices

LEA-6 is a medium-sized district in central California, and it joined the study later than other LEAs. One elementary school, one middle school, and one high school collaborated with HumRRO and participated in data collection activities. HumRRO conducted the site visit to LEA-6 in January of 2019. Just under half of LEA-6's student population is classified as socioeconomically disadvantaged, which is less than the state average (62%), and participating schools have very few English learners (1–2%). According to 2018 test results, close to half the students in the study schools met or exceeded grade level academic achievement standards in ELA, but fewer did so for mathematics (14%–35%). PLCs are not established across the LEA, though the middle school created mathematics PLCs to address the low mathematics performance. The PLCs met to identify a learning objectives pathway for each grade and found students have significant knowledge gaps. The elementary school has no PLC but holds twice-

monthly meetings to share information, and the high school has five days a year for professional development but not ongoing release time for collaboration.

LEA-6 teachers reported inadequate student access to and familiarity with technology, especially with respect to the CAASPP testing platform. LEA-6 uses Chromebooks, and the ratio of computers to students has been narrowing but is estimated to be about one-to-three. Teachers find scheduling shared use of the resources challenging, and many students do not have computers at home. Supplemental programs (e.g., Freckle, Typing Agent, Accelerated Reader) purchased at the elementary school give younger students opportunities to use technology, which has increased student motivation.

Regarding CAASPP summative data from 2018, LEA-6 downloaded results into Illuminate (formerly SchoolCity) and Aeries, and analyzed results to help prioritize which professional learning to offer teachers. Principals reviewed the data together, and discussed the highlights, and areas for improvement. The principals were then responsible for distributing data to teachers, who discussed scores by grade levels. Though previous years' scores and other reports can be accessed through Illuminate, not all teachers or schools are aware of this. At the middle school, teachers compared school and classroom results to state averages; their findings inspired improvements through a mixture of increasing student motivation, ensuring an appropriate testing environment, and improving instructional practice.

LEA-6 strongly encouraged IAB use but was careful to avoid mandates and specificity in how schools implement use. The LEA conveyed the idea that IABs are useful in exposing students and teachers to expectations and test item types for Smarter Balanced summative assessments in ELA and mathematics, and to inform teachers of instructional gaps. Across the study schools, students in high school had the most opportunities (21) to take ELA IABs (10 different tests), and students in grade eight had the most opportunities (18) to take mathematics IABs (5 different tests). Most IAB's were given in the standardized manner; however, some IABs were given in nonstandardized manner to allow students to work in teams to review difficult concepts. Also, some IABs were administered in science and history high school classes to get "buy in" from teachers of other content areas. In addition to using IABs to familiarize students with the summative format and content, teachers used IABs to identify instructional gaps, set grade-level expectations, monitor progress, and improve student motivation.

LEA-6 identified several concerns about the operations of the Smarter Balanced assessments (e.g., awkward pausing in the text-to-speech feature, inability to delete results from IABs given only to test bandwidth).

There was considerable consensus in LEA-6 that the DL has improved since it was first rolled out and that access, while still difficult, has gotten easier. Fewer teachers use the DL than do not. Many teachers would like to learn how to link IABs to the DL; however, time for training is limited and the DL competes for attention with other resources.

HumRRO identified the following best practices supporting effective use of CAASPP components to improve teaching and student learning:

- Allowing school- and teacher-level flexibility increased the use of IABs. LEA-6 learned from previous years that mandating how and what IABs should be given at the LEA-level resulted in teacher frustration.
- Allowing local decision-making enhanced buy-in and allowed teachers to integrate IABs that fit with their instruction practices.
- In schools where teachers have time to collaborate, share scores and analyze assessment results in teams, use of IABs has led to more positive attitudes towards CAASPP testing, using IABs, and the DL.

LEA-7 Summary of Findings and Best Practices

LEA-7 is a medium-sized school district in central California. LEA-7 joined the study in April 2019 and schools were available for only a limited number of data gathering events. Three elementary schools and grades six through eight in one K–8 school participated, with the latter representing middle school CAASPP use in the district. The study schools have high majority socioeconomically disadvantaged populations (78%–96%) and a significant percentage of ELs (28%–38%). To address its large population of new students from other countries, LEA-7 developed a newcomer program, which gives students an opportunity for intensive English language learning to help them integrate into regular school activities. Regarding 2018 academic achievement, lower percentages of students at the elementary schools met or exceeded grade level standards in ELA (21%–30%) or mathematics (13%–22%) compared to students at the middle school (ELA, 55%; mathematics, 39%). In the study schools, teachers within each grade level function as a PLC.

LEA-7 increased the number of devices to achieve one-to-one computing programs for students in grades three through eight. Elementary schools also provide typing practice devices and programs for grade three.

LEA-7 downloaded students' 2018 CAASPP summative data from ORS and uploaded them to Illuminate, to avoid what they see as a cumbersome rostering process. Housing the data in Illuminate also allows for other assessments' data to be in the same system, enabling teachers to access multiple measures for each student. Schools reviewed data during teacher meetings, and some teachers reviewed results with students to help them establish goals for performance on the 2019 Smarter Balanced summative assessments in ELA and mathematics. The timing of goal setting varied from the beginning of the year to the weeks leading up to the summative assessment. The goal setting process was intended to motivate students to greater levels of achievement. One elementary school reviewed five-by-five placement reports, available on CDE's Dashboard website, which combines data from the current and prior years to produce a color-coded grid with five Status levels and five Change levels.

LEA-7 mandated that schools administer one IAB in ELA and one in mathematics during 2018–19. It encouraged schools to administer more IABs, indicating data showed higher IAB use was associated with individual scale score growth. An administrator at one elementary school established a schedule for each grade indicating which IAB

should be given in which week. Other schools left these decisions to grade-level PLCs or individual teachers. Across the study schools, students in grade three had the most opportunities (44) to take ELA IABs (9 different tests) and the most opportunities (43) to take mathematics IABs (6 different tests). Schools administered IABs in standardized and nonstandardized manners. In addition to using IABs to familiarize students with the summative format and tools, teachers at the study schools used results from the assessment to guide instructional decisions. For example, teachers reviewed results and constructed small groups to provide differentiated instruction according to student needs. Other teachers conducted item analyses to identify particularly difficult items and then developed whole group instruction around the concepts and strategies necessary for addressing these items. Teachers also reported using the results to assess the effectiveness of their instruction. Although the school's CAASPP coordinator provided teacher-specific data, teachers would like more direct access to data and item-by-item results. There are ways to create rosters of students within the IA Reporting System; however, this school seems unaware of that functionality.

None of the schools' teachers reported using DL resources. Teachers suggested they had many other resources (e.g., Benchmark, Eureka Math, Newsela, Prodigy, Zearn) to help with lesson planning and formative assessment.

HumRRO identified the following best practices supporting effective use of CAASPP components to improve teaching and student learning:

- The LEA enabled teachers to be greater consumers of assessment data by (a) uploading CAASPP results into an existing LEA data system that includes other data sources and (b) training teachers on how to access and use these data.

- One elementary school used available funding to conduct Saturday boot camps that trained teachers on use of Smarter Balanced summative assessment data and IABs. Teachers who engaged in this training were much more effective in their use of IABs for students.

- Teachers made a focused effort to orient students to the Smarter Balanced computer interface using IABs while also using the IABs to understand Smarter Balanced item types on the summative assessment and in some cases guide instructional decision making.

This page is intentionally blank.

Appendix E: CAST Alignment Workshop Panelist Training

The content of this appendix was presented as large group training to CAST Alignment Workshop panelists on February 28, 2019. The format of the training was oral delivery with a slide presentation. The text below provides all content from the presentation slides.

Workshop Logistics

Welcome

- Introductions
  - HumRRO staff
  - California Department of Education staff
- Participants list
- Panelist recruitment – different sources from across California, all with CA NGSS expertise
  - Science Subcommittee of the Curriculum and Instruction Committee
  - NGSS Early Implementers Committee
  - Response to CAASPP Update invitation for Science educators
- Panel assignments

Meeting Agenda and Goals

- Large group training
  - Overview of California NGSS (CA NGSS)
  - Overview of alignment method
- Small group activity – Iterative sessions
  - Independently review small number of CAST items online and make ratings in Excel sheet.
  - Discuss ratings as a panel.
  - Review metadata as a panel.
  - Final discussion and consensus ratings.
- Outcomes of workshop
  - Contribute to HumRRO's standalone CAST Alignment Study Report and 2020 Independent Evaluation Report.
  - Enhance panelists' working knowledge of CA NGSS, the CAST assessment, and alignment methods.

Panelist Large Group Training

Looking Ahead for Today and Beyond…

- Who is HumRRO?
- Roles:
  - HumRRO staff…provide alignment expertise, facilitation, and final research report
  - Panelists…expert reviewers provide the data used to analyze and report alignment findings
  - CDE staff…observers and CAST experts who will monitor progress of HumRRO, answer panelists' questions, and be available as a resource.

Alignment Basics and Process

Define concept of alignment.

- Explain why alignment is important.
- Understand how alignment is measured.
- Understand alignment-related concepts.
- Participate in specific training on cognitive complexity ratings.

Alignment Basics (1)

What is Alignment?

> "The degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what is expected."
>
> -Webb, 2005

Alignment Basics (2)

Why is Alignment Important?

- Fairness for all students…
  - Consistency in general curriculum
  - Accurate assessment of what students can do and are expected to know from State content standards and the curriculum
  - Improves teacher instruction and student learning.
- Federal statutes require alignment.

Alignment Basics (3)

How is Alignment Measured?

- Alignment is a matter of degree, not all or none.
- Science, particularly NGSS three-dimensional standards, stymies traditional alignment methods.
    1) There is no such thing as a "standard." Standards exist as interactions among dimensions and content domains. This means that items may not measure one identifiable "thing."
    2) Depth-of-Knowledge or cognitive complexity of items is not readily matched to expectations from the standards.
    3) Balancing dimension and domain requirements is more challenging than simply addressing groups of reporting standards.
- So, what are we going to do?

Science Alignment Training

Alignment: what does that mean for a CAST item?

- Overview of CA-NGSS
- Overview of DOK for application to science items
- Group Discussion of CA-NGSS
- Practice with DOK
- Overview of Alignment Methodology
- Overview of Rating Process

The Alignment Rule

For items, "aligned" means the item's content is included in the standard (Performance Expectation (PE)—may be a DCI, SEP, or CCC or some combination).

- No item can capture the full breadth of a PE! Think of the PE (or DCI, SEP, or CCC) as buckets. If the item fits in any of these buckets, it is aligned. It does not need to fill the bucket.
- There is some overlap within the dimensions. We want to identify the main or primary dimension—and a secondary if it is clearly indicated. Vagueness within the dimensions does NOT indicate a secondary match.

Applying the Alignment Rule

- Panelists will review CAST items and identify the specific PEs and dimensions (SEP, DCI, CCC).
- HumRRO will analyze panelists' ratings to answer, “What proportion of items address two or more dimensions?”

Overview of CA NGSS

Three Dimensional Learning

Dimension 1: Scientific and Engineering Practices (SEP)
Dimension 2: Crosscutting Concepts (CCC)
Dimension 3: Disciplinary Core Ideas (DCI)

CA NGSS Structure

5-PS2 Motion and Stability: Forces and Interactions

Students who demonstrate understanding can:

5-PS2-1 **Support an argument that the gravitational force exerted by Earth on objects is directed down.** [Clarification Statement: "Down" is a local description of the direction that points toward the center of the spherical Earth.] [Assessment Boundary: Assessment does not include mathematical representation of gravitational force.]

The performance expectations above were developed using the following elements from the NRC document *A Framework for K–12 Science Education*:

Science and Engineering Practices

Engaging in Argument from Evidence

Engaging in argument from evidence in 3–5 builds on K–2 experiences and progresses to critiquing the scientific explanations or solutions proposed by peers by citing relevant evidence about the natural and designed world(s).

- Support an argument with evidence, data, or a model. (5-PS2-1)

Disciplinary Core Ideas

PS2.B: Types of Interactions

- The gravitational force of Earth acting on an object near Earth's surface pulls that object toward the planet's center. (5-PS2-1)

Crosscutting Concepts

Cause and Effect

- Cause and effect relationships are routinely identified and used to explain change. (5-PS2-1)

Connections to other DCIs in fifth grade: NA

Articulation of DCIs across grade-bands: **3.PS2.A** (5-PS2-1); **3.PS2.B** (5-PS2-1); **MS.PS2.B** (5-PS2-1)**; MS.ESS1.B** (5-PS2-1); **MS.ESS2.C** (5-PS2-1)

California Common Core State Standards Connections:

ELA/Literacy –

RI.5.1 Quote accurately from a text when explaining what the text says explicitly and when drawing inferences from the text. (5-PS2-1)

RI.5.9.a,b Integrate information from several texts on the same topic in order to write or speak about the subject knowledgeably. (5-PS2-1)

W.5.1.a–d Write opinion pieces on topics or texts, supporting a point of view with reasons and information. (5-PS2-1)

Interaction Between CA NGSS and Item Specifications

Alignment study links items to CA NGSS, not to item specifications—as required by federal peer review.

- Item specifications provide additional context for how the standards document was interpreted and used to generate items.
- It is important to consider both when conducting an alignment study. The item specifications provide a window into the intentions of the assessment program.
- Match process for items—Identify PE (from CA NGSS), find associated item specifications for that PE (in Item Specs), Identify DCI (code + number + letter), SEP, and CCC. Provide any notes or comments.

Overview of DOK

Introduction (1)

- Overview
  - Webb's criteria for aligning standards and standardized assessments
  - Curricular elements can be categorized by cognitive demands.
  - Categories reflect cognitive expectation.
  - Measure of cognitive complexity NOT difficulty

Introduction (2)

- Level of Cognitive Complexity
  - Recall and Reproduction (DOK1)
  - Working with Skills and Concepts (DOK2)
  - Short-term Strategic Thinking (DOK3)
  - Extended Strategic Thinking (DOK4)
- CA NGSS
  - DOKs might not fully reflect the range of higher-order thinking of 3D standards

DOK1 – Recall and Reproduction (1)

- Basic tasks
  - Recall or reproduce
- Knowledge and/or skills
- Content
  - Facts
  - Terms

- Properties of objects
- Application of simple procedures and/or formulas

DOK1 – Recall and Reproduction (2)

- Restrictions
  - Minimal transformation
  - Nominal extended processing
- Requires low cognitive demand
- Key tasks
  - List
  - Identify
  - Define
- Higher DOK items may contain parts that can be classified as a lower DOK

DOK1 – Recall and Reproduction (3)

Example:

Which of the following IS NOT a variable contributing to a successful hunter-gatherer community?

A. Water
B. Edible plants
C. Hospitable climate
D. Well-paved roads

http://thinkingschoolsacademy.org/wp-content/uploads/2014/03/Creating-Tests-for-Higher-Order-Thinking.pdf

DOK2 – Skills and Concepts (1)

- Engagement of some mental processing
  - Compare or contrast
- People, places, events, and concepts
  - Convert information
  - Classification into meaningful categories
  - Describe or explain
- Issues, problems, patterns, relationships… etc.

DOK2 – Skills and Concepts (2)

- Applications
  - Laboratory setting
  - Content
- Principles, categories, heuristics, and protocols
- Key tasks
  - Summarize
  - Infer
  - Classify
  - Estimate

DOK2 – Skills and Concepts (3)

Example:

Look at the two magnets above. If you push the two magnets toward each other as shown, the magnets will:

A. Break into many pieces.
B. Turn in opposite directions.
C. Be pushed away from each other.
D. Be pulled toward each other.

http://www.e-archives.ky.gov/pubs/Education/cca_science_support.pdf

DOK3 – Strategic Thinking (1)

- Higher order thinking processes
  - Analysis and evaluation
- Solve real-world problems with predictable outcomes
  - Rationale
  - Solution in a project-based setting
  - Key tasks
- Analyze, explain and support with evidence, generalize, and create

DOK3 – Strategic Thinking (2)

Example:

The typical spring and summer has resulted in much less rainfall than normal. This will most likely result in:

A. A hot summer
B. Fewer healthy plants and vegetables to harvest
C. A large harvest of edible plants and fruit
D. Trees which have grown faster than normal

http://thinkingschoolsacademy.org/wp-content/uploads/2014/03/Creating-Tests-for-Higher-Order-Thinking.pdf

DOK4 – Extended Thinking (1)

- Extended use of higher order thinking
  - Synthesis, reflection, assessment and adjustment of plans over time
- Solve real-world problems with unpredictable outcomes
  - Employing and sustaining strategic thinking over a longer period of time
  - Key tasks
- Synthesize, reflect, conduct, and manage

DOK4 – Extended Thinking (2)

Example:

Analyze the performance of a vehicle that uses the conservation of momentum to propel itself. Design a method to improve the efficiency of this process.

Detailed Descriptors of Depth-of-Knowledge Levels for Science

Level 1 - Recall & Reproduction

a. Recall or recognize a fact, term, definition, simple procedure (such as one step), or property
b. Demonstrate a rote response
c. Use a well-known formula
d. Represent in words or diagrams a scientific concept or relationship
e. Provide or recognize a standard scientific representation for simple phenomenon
f. Perform a routine procedure, such as measuring length
g. Perform a **simple** science process or a set procedure (like a recipe)
h. Perform a clearly defined set of steps
i. Identify, calculate, or measure

***Note*: If the knowledge necessary to answer an item automatically provides the answer, it is a Level 1.**

Level 2 - Skills & Concepts

a. Specify and explain the relationship between facts, terms, properties, or variables
b. Describe and explain examples and non-examples of science concepts
c. Select a procedure according to specified criteria and perform it
d. Formulate a routine problem given data and conditions
e. Organize, represent, and compare data
f. Make a decision as to how to approach the problem
g. Classify, organize, or estimate
h. Compare data
i. Make observations
j. Interpret information from a simple graph
k. Collect and display data

***Note*: If the knowledge necessary to answer an item does not automatically provide the answer, then the item is at least a Level 2. Most actions imply more than one step.**

Level 3 - Strategic Thinking

a. Interpret information from a complex graph (such as determining features of the graph or aggregating data in the graph)
b. Use reasoning, planning, and evidence
c. Explain thinking (beyond a simple explanation or using only a word or two to respond)
d. Justify a response
e. Identify research questions and design investigations for a scientific problem
f. Use concepts to solve non-routine problems/more than one possible answer
g. Develop a scientific model for a complex situation
h. Form conclusions from experimental or observational data
i. Complete a multi-step problem that involves planning and reasoning
j. Provide an explanation of a principle
k. Justify a response when more than one answer is possible
l. Cite evidence and develop a logical argument for concepts
m. Conduct a designed investigation
n. Research and explain a scientific concept
o. Explain phenomena in terms of concepts

***Note*: Level 3 is complex and abstract. If more than one response is possible, it is at least a Level 3 and calls for use of reasoning, justification, evidence, as support for the response.**

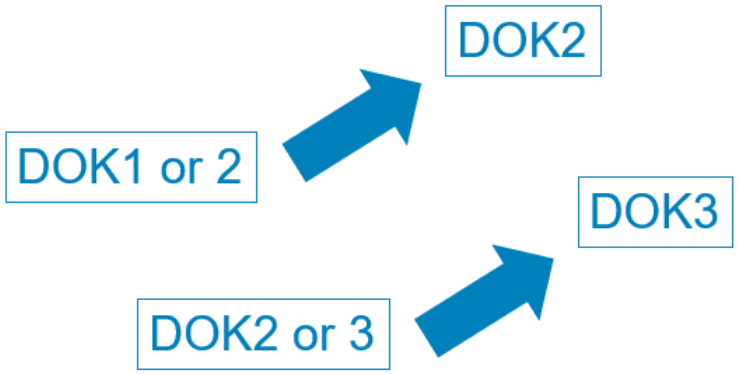
Level 4 - Extended Thinking

a. Select or devise approach among many alternatives to solve problem
b. Based on provided data from a complex experiment that is novel to the student, deduct the fundamental relationship between several controlled variables.
c. Conduct an investigation, from specifying a problem to designing and carrying out an experiment, to analyzing its data and forming conclusions
d. Relate ideas *within* the content area or *among* content areas
e. Develop generalizations of the results obtained and the strategies used and apply them to new problem situations

***Note*: Level 4 activities often require an extended period of time for carrying out multiple steps; however, time alone is not a distinguishing factor if skills and concepts are simply repetitive over time.**

Source: K. Hess, Center for Assessment, based on Webb, update 2005

DOK Selection



Discussion

Appropriateness of DOK NGSS
Any questions?

Overview of Alignment Methodology

ACHIEVE Criteria

High-quality, aligned summative science assessments:

a) Focus on performance expectations that form the core of scientific concepts and skills described in the standards;
b) Assess students' ability to use scientific and engineering practices to understand specific, disciplinary core ideas and apply their understanding of crosscutting concepts;
c) Assess a balance of physical sciences, life sciences, Earth and space sciences;
d) Require a range of cognitive demand; and
e) Ensure high-quality items and a variety of item types.

Item Ratings

Depth of Knowledge of item

- Alignment to the identified CA NGSS standard (PE)
- Does alignment require student to demonstrate grade-appropriate understanding of and facility with…
  - SEP?
  - DCI?
  - CCC?

- Are there any issues with the quality of the item or phenomena?
- Other issues or notes

Ratings Process

In grade-level panels, small group training on rating spreadsheet and online access to CAST items

- Group calibration (1-3 items)
- Independent ratings (5-10 items per session)
- Group consensus (by session)
- Discussion of notes and item issues (e.g., phenomena, item quality, etc.)
- Whole test debriefing
- Workshop evaluation

Group Discussion of CA NGSS

Guiding Questions

What if there is overlap between two PEs, and I can't tell which PE is the most appropriate for an item?

- The SEPs and CCCs can be vague. Where does one stop and the other begin (e.g., analyzing and interpreting data versus using mathematical and computational thinking)?
- The DCI, SEP, and CCC are linked to specific PEs. Should we expect items to be related to linked dimensions?
- What if an item addresses a fairly minor point or does not capture the "heart" of the PE?
- What if the phenomenon the item is based on is poor, contrived, or just not appropriate for the item?
- Other questions or topics?

Questions?

Sample Item #1 Grade 8

Ethan is observing chemical and physical properties of a substance. He heats a substance and observes that the substance turns from a brown solid to a black powder. He refers to several chemistry journals that claim this represents a chemical reaction. From his observation and research, he concludes that the substance goes through a chemical change when heated. How can Ethan **best** defend his conclusion?

A. by demonstrating that the substance will eventually melt if the temperature continues to increase

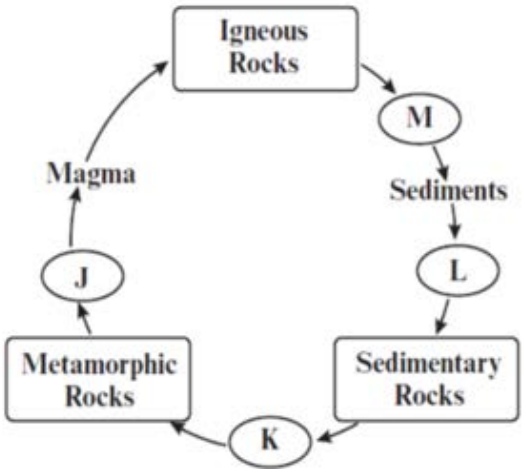
B. by verifying that the substance is now made up of different molecules than before it was heated

C. by verifying that the substance is made up of only one type of element

D. by demonstrating that the substance is less dense after it is heated

Sample Item #2 Grade 8

Ice forms in the cracks of a basalt rock formation and breaks some rock into smaller pieces. The diagram below shows part of the rock cycle.



At which point in the cycle shown above would the process of breaking down rocks occur?

A. J

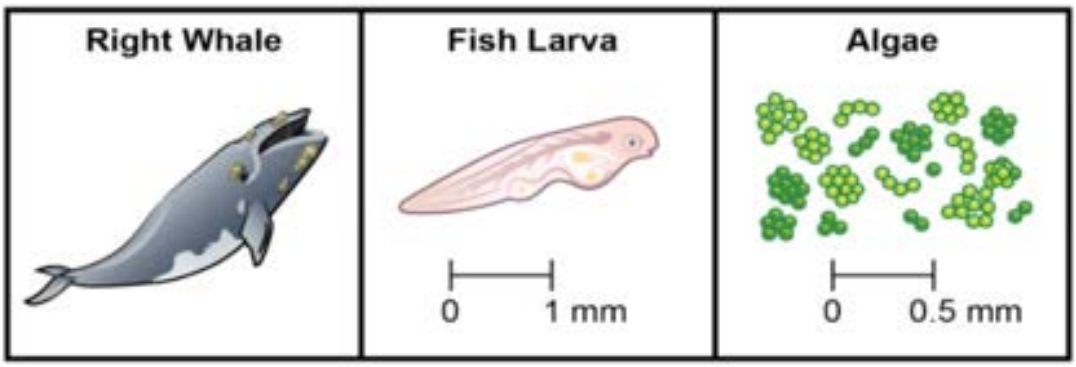
B. K

C. L

D. M

Sample Item #3 Grade 5

A group of scientists is studying organisms in an ocean ecosystem. They show you three pictures of the organisms they see.



- Right whales grow to be about 15 meters (m) long. That is about the length of a tractor-trailer.
- Fish larva are young fish that just hatched from an egg. They are only a few millimeters (mm) in length, which is smaller than a sesame seed
- Algae are plants that live in the ocean. They are smaller than the period at the end of this sentence.

Scientists' Observations: Flow of Energy and Matter

The scientists watch the right whales eating the fish larva as they swim in the ocean. They also watch the tiny fish larva eating algae as they swim in the ocean. More fish larvae are found in areas of the ocean that have more algae.

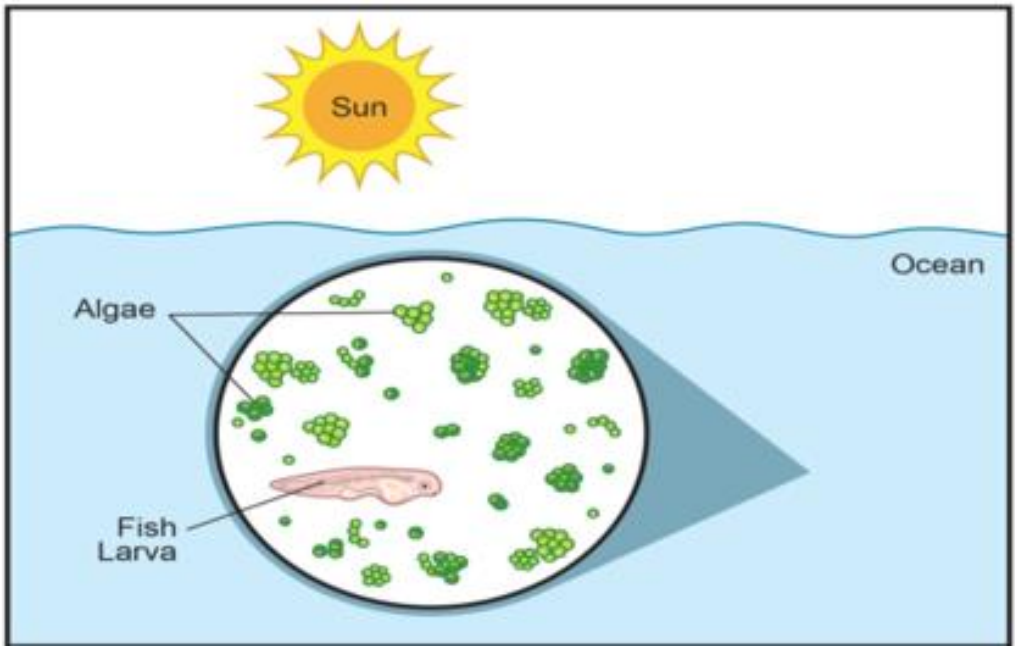
Question 1. Which statements explain what happens when right whales eat the fish larva as they swim in the ocean? Select the three correct answers.

1. Whales get matter they need to grow.
2. Whales get energy they need to swim.
3. Energy is transferred from the whales to the fish larva as the whales eat.
4. Energy and matter are transferred from the fish larva to the whales as the whales eat.
5. Matter is transferred from the water to the whales and the fish larva as the whales eat.

Science Example #3 Grade 5 (cont.)

Question 2. The scientists want you to model the flow of energy and matter through the ecosystem as fish larva eat the algae. The scientists have a diagram showing the Sun, fish larva, and algae.

Complete the model by drawing **two** arrows to show how energy flows among the Sun, the fish larva, and the algae. The arrows should point in the direction energy flows.



Appendix F:
Detailed Descriptions of Figures with Images

Figure 1.1. Framework for an integrated set of assessments (p. 1-6)

- The figure is composed of two diagrams side by side.
- The diagram on the left is a picture of a car wheel overlaid with five text circles to represent the cycle of improving teaching and learning. The circle in the center of the wheel is labeled Improving Teaching and Learning. The other four text circles are positioned around the tire (top, right, bottom, and left). The top circle is labeled Instruction. Moving clockwise, the next circle (right) is Curriculum, followed by Assessment (bottom) and Curriculum (left).
- The diagram on the right has a similar design to the car wheel but includes just a center text circle with four surrounding text circles (top, right, bottom, left). The center circle is labeled Improving Teaching and Learning. The top circle is labeled Summative Results with an arrow pointing clockwise to the second circle (right) labeled Digital Library. Double-headed arrows (indicating going back and forth) link the second to the third circle (bottom), labeled Interim Results. Double-headed arrows indicate back and forth between the third (bottom) and fourth (left) circle, which is also labeled Digital Library. There is a single arrow leading from the fourth circle (left) back to the first circle (top).

Figure 2.1. Hierarchy of Smarter Balanced item development (p. 2-11)

- The figure is comprised of 13 text boxes. There is one text box at the left (centered vertically) and to its right are three sets of four vertically arranged text boxes.
- The text in the box on the left is Overall Claim (Content Domain for ELA or Math). Four lines connect this box to each of the four boxes in the first set of vertical boxes.
- In the first set of vertical boxes, the text in the top box is Claim 1 (Sub-Domain). The text in the box below is Claim 2 (Sub-Domain) and the same pattern continues for the next two boxes (Claims 3 and 4). A horizontal line connects each of the Claim boxes to a box in the next set of vertical boxes to the right.
- The text in every box in the second set of vertical boxes is Assessment Target(s). Again, horizontal lines connect each Assessment Target(s) box to a box in the next set of vertical boxes to the right.
- The text in every box in the third (rightmost) set of vertical boxes is Standard(s).