2016 No. 077

California Assessment of Student Performance and Progress (CAASPP): 2016 Independent Evaluation Report Volume 1

Prepared for:	California Department of Education Assessment Development and Administration Division 1430 N Street, Suite 4401 Sacramento, CA 95814-5901	**Prepared under:**	CN140472.1
Lead Authors:	Michele M. Hardoin Lauress Wise Arthur Thacker Emily Dickinson Caroline Wiley	**Date:**	November 21, 2016
Editors:	Christa Watters D.E. (Sunny) Becker Sheila Schultz		

Headquarters: 66 Canal Center Plaza, Suite 700, Alexandria, VA 22314 | Phone: 703.549.3611 | Fax: 703.549.9025 | humrro.org

Table of Contents

Executive Summary ES-1
- Major Findings and Recommendation from the Smarter Balanced Interim Assessment Administration Study ES-1
- Major Findings and Recommendation from the Access to Designated Supports and Accommodations Study ES-4
- Major Finding and Recommendation from Observations of CAASPP Training Sessions ES-6
- General Summary ES-7

Chapter 1: Introduction 1-1
- Background 1-1
- Summary of 2015 Evaluation Activities 1-2
- Organization and Contents of the CAASPP 2016 Evaluation Report 1-4

Chapter 2: Implementing the Independent Evaluation Plan 2-7
- Background Research on CAASPP System 2-7
- Safeguarding Confidential Data 2-7
- Observation of CAASPP Educator Training Sessions 2-8
- Creation of a Collaborative Research Network with Local Educational Agencies 2-9

Chapter 3: CAASPP Smarter Balanced Interim Assessment Administration Study 3-15
- Background 3-15
- Study Design 3-17
- Focus Groups and Interviews 3-21
- Statewide Survey Methods 3-27
- Statewide Survey Findings 3-36
- Analysis of Statewide Testing Data 3-82

Chapter 4: Access to Designated Supports and Accommodations Study 4-84
- Background 4-84
- Study Design 4-86
- Methods and Analysis 4-87

Chapter 5: Update to the CAASPP Independent Evaluation Study Plan 5-103
- Original Independent Evaluation Study Plan 5-103
- Responding to Revised Timelines for New CAASPP Assessments 5-104
- Summary of 2016 Updates to the Independent Evaluation Studies 5-104

Chapter 6: Findings and Recommendations 6-107
- Results from Smarter Balanced Interim Assessment Administration Study (Chapter 3) 6-107
- Results from Access to Designated Supports and Accommodations Study (Chapter 4) 6-120
- Results from Observations of CAASPP Training Sessions (Chapter 2) 6-129

References 133

Glossary of Acronyms 135

Table of Contents (Continued)

List of Tables

Table 2.1. Classification of 2015–16 LEARN Members by Selection Criteria 2-12

Table 3.1. Research Question Coverage by Data Source 3-17
Table 3.2. Interim Assessment Administration Study Data Source Descriptions 3-19
Table 3.3. Smarter Balanced Interim Assessment Study LEA Focus Group and Interview Samples 3-21
Table 3.4. Summary of Common Themes in Interim Assessment Study 3-23
Table 3.5. Sampling Frame Descriptions for Each Survey 3-28
Table 3.6. Survey Response Rates and Respondent Behaviors 3-33
Table 3.7. Survey Exclusion Criteria by Survey Path 3-34
Table 3.8. Survey LEA Demographics Compared to the State 3-35
Table 3.9. Descriptive Statistics for School Site Coordinator and Test Administrator Analytic Samples 3-36
Table 3.10. Descriptive Statistics for LEA Coordinator Analytic Sample 3-36
Table 3.11. Survey Background Questions and Appendix Tables Cross-Reference 3-37
Table 3.12. Most Commonly Reported Titles for Each Survey 3-37
Table 3.13. Interim Assessment Usage Survey Questions and Appendix Tables Cross-Reference 3-38
Table 3.14. Count of Interim Assessments Administered During 2014–15 and 2015–16 3-40
Table 3.15. Reported Interim Assessment Policies 3-41
Table 3.16. Percentage of Test Administrators Reporting Interim Assessments as a Requirement Matched by LEA 3-42
Table 3.17. Most Commonly Reported Combinations of Interim Assessment Type Usage 3-43
Table 3.18. Interim Assessment Administration Settings Reported by TAs 3-45
Table 3.19. Interim Assessment Preparation and Training Survey Questions and Appendix Tables Cross-Reference 3-46
Table 3.20. Interim Assessment Impacts on Instruction Survey Questions and Appendix Tables Cross-Reference 3-54
Table 3.21. Interim Assessment Administration Survey Questions and Appendix Tables Cross-Reference 3-60
Table 3.22. Reported Standardization of Administration Procedures by School Site Coordinators 3-60
Table 3.23. Reported Standardization of Administration Procedures by Test Administrators 3-61
Table 3.24. Interim Assessment Accommodations Survey Questions and Appendix Tables Cross-Reference 3-62
Table 3.25. Interim Assessment Challenges Survey Questions and Appendix Tables Cross-Reference 3-65

Table of Contents (Continued)

List of Tables

Table 3.26. Interim Assessment Improvement Survey Questions and Appendix Tables Cross-Reference 3-74
Table 3.27. Frequencies of Non-Interim Assessment Users by Survey 3-77
Table 3.28. Non-Use Survey Questions and Appendix Tables Cross-Reference 3-77
Table 3.29. TA Primary Factors for Not Administering Interim Assessments 3-78
Table 3.30. LEA Primary Factors for Not Administering Interim Assessments 3-79
Table 3.31. SITE Primary Factors for Not Administering Interim Assessments 3-79
Table 3.32. LEA Primary Reasons for Allowing Interim Assessment Use in 2016–17, of Those That Did Not Use Them in 2015–16 3-80
Table 3.33. SITE Primary Reasons for Allowing Interim Assessment Use Next Year, of Those That Did Not Use Them in 2015–16 3-81
Table 3.34. Interim Assessment Usage and Mean Scale Score (2015–16) 3-83
Table 3.35. Interim Assessment Usage and Mean Scale Score Change (2015–2016) 3-83

Table 4.1. Characteristics of LEAs Participating in the Access to Designated Supports and Accommodations Study 4-88
Table 4.2. Characteristics of Participating Schools 4-89
Table 4.3. Planned Sample vs Research Sample 4-89
Table 4.4. Number of Site Visit Components Completed 4-90
Table 4.5. Most Frequently Reported Instructional and Classroom Testing Supports/Accommodations 4-94
Table 4.6. Interim Assessment Accommodations Survey Questions and Appendix Tables Cross-Reference 4-95
Table 4.7. Percentage of Students Taking 2015–16 Smarter Balanced Assessments Identified as SWD or EL 4-98
Table 4.8. Percentage of Students Offered Accommodations, Designated Supports, and Universal Tools 4-99
Table 4.9 Students with Matched Checklist and Summative Assessment Data 4-99
Table 4.10. Match Between Instructional Practices and Summative Assessment Accommodations, Designated Supports, and Universal Tools 4-100
Table 4.11. Average Number of Accommodations and Designated Supports Offered to Students Statewide 4-101

Table of Contents (Continued)

List of Figures

Figure 3.1. Illustration of Smarter Balanced Assessment System. 3-15
Figure 3.2. Staged design of Interim Assessment Administration Study. 3-18
Figure 3.3. LEA CAASPP coordinator survey blueprint. 3-29
Figure 3.4. School site CAASPP coordinator survey blueprint. 3-30
Figure 3.5. Test administrator survey blueprint. 3-31
Figure 3.6. General path for each role (survey). 3-39
Figure 3.7. Reported usage of interim assessments (LEA and SITE surveys). 3-40
Figure 3.8. Primary uses that contributed to requiring or highly encouraging interim assessments (LEA and SITE surveys). 3-44
Figure 3.9. Third-party assessment usage (LEA and SITE surveys). 3-46
Figure 3.10. Helpfulness of Smarter Balanced Interim Assessment resources (LEA survey). 3-47
Figure 3.11. Ease of Smarter Balanced Interim Assessment Systems (LEA survey). 3-49
Figure 3.12. Helpfulness of Smarter Balanced Interim Assessment Resources (SITE survey). 3-50
Figure 3.13. Ease of Smarter Balanced Interim Assessment Reporting System (SITE survey). 3-51
Figure 3.14. Helpfulness of Smarter Balanced Interim Assessment resources (TA survey). 3-52
Figure 3.15. Ease of Smarter Balanced Interim Assessment Systems (TA survey). 3-53
Figure 3.16. Ease of Smarter Balanced Interim Assessment Reporting System (TA survey). 3-54
Figure 3.17. Usefulness of Smarter Balanced Interim Assessments (LEA survey). 3-55
Figure 3.18. Decisions about educator supports and students (LEA survey). 3-56
Figure 3.19. Usefulness of Smarter Balanced Interim Assessments (SITE survey). 3-57
Figure 3.20. Decisions about students and educator support (SITE survey). 3-57
Figure 3.21. Usefulness of Smarter Balanced Interim Assessments (TA survey). 3-58
Figure 3.22. Decisions about students (TA survey). 3-59
Figure 3.23. Confidence in the validity of results (LEA, SITE, and TA surveys). 3-59
Figure 3.24. Fidelity to test administration procedures (TA survey). 3-61
Figure 3.25. Accommodations for students with disabilities and supports for English learners (SITE survey). 3-63
Figure 3.26. Accommodations for students with disabilities and supports for English learners (TA survey). 3-64
Figure 3.27. Challenge of administration (LEA survey). 3-65
Figure 3.28. Challenge of test content (LEA survey). 3-66

Table of Contents (Continued)

List of Figures

Figure 3.29. Challenge of hand scoring (LEA survey). 3-67
Figure 3.30. Challenge of reporting (LEA survey). 3-68
Figure 3.31. Challenge of administration (SITE survey). 3-69
Figure 3.32. Challenge of test content (SITE survey). 3-69
Figure 3.33. Challenge of hand scoring (SITE survey). 3-70
Figure 3.34. Challenge of reporting (SITE survey). 3-71
Figure 3.35. Challenge of administration (TA survey). 3-72
Figure 3.36. Challenge of test content (TA survey). 3-72
Figure 3.37. Challenge of hand scoring (TA survey). 3-73
Figure 3.38. Challenge of reporting (TA survey). 3-74
Figure 3.39. Most important changes for 2016–17 (LEA and SITE surveys). 3-75

Figure 4.1. Accommodations for students with disabilities and supports for English learners (SITE survey) 4-96
Figure 4.2. Accommodations for students with disabilities and supports for English learners (TA survey). 4-97

This page is intentionally blank.

EXECUTIVE SUMMARY

Pursuant to California *Education Code* (*EC*) Section 60649, the Human Resources Research Organization (HumRRO) is conducting a three-year independent evaluation for the California Assessment of Student Performance and Progress (CAASPP). The purpose of the evaluation is to provide objective technical advice and consultation on activities supporting the implementation of the CAASPP System. The Independent Evaluation Study Plan was approved by the State Board of Education in September 2015.

This annual report covers the following key activities HumRRO conducted as the independent evaluator during the 2015–16 school year:

- Implementation of the Independent Evaluation Study Plan
- Smarter Balanced Interim Assessment Administration Study
- Access to Designated Supports and Accommodations Study
- Update to the CAASPP Independent Evaluation Study Plan

The multiple systems that form the new California assessment environment are complex. Across the state, local educational agencies (LEAs), schools, and teachers are learning to implement the various components of the CAASPP system, while those very components are being refined and improved by the CDE, its testing contractors, and the Smarter Balanced Assessment Consortium. The transitions required to implement the new system include delivering instruction aligned to the new academic standards[1], delivering assessments using a new online platform, learning new procedures to locally hand score constructed response items for interim assessments, applying new knowledge about the interconnectedness of the technical aspects of the various CAASPP System components and supports, and interpreting new forms of assessment results.

Our recommendations based on the 2015–16 evaluation activities acknowledge that this major transition is still in its early stages. Future evaluation reports will incorporate activities designed to monitor the system as it matures.

Major Findings and Recommendation from the Smarter Balanced Interim Assessment Administration Study

The Smarter Balanced Interim Assessment Administration Study was designed to explore: (a) how LEAs decided whether and which interim assessments to administer in 2015–16, (b) how smoothly the administration and scoring of these assessments went, and (c) the extent to which the results may have had an impact on instruction. LEAs have the authority to decide whether to use interim assessments provided as part of the Smarter Balanced Assessment System, which include Interim Comprehensive Assessments and Interim Assessment Blocks.

The data collection methods for the study consisted of a small set of focus groups followed by three online surveys administered to a sample of (a) LEA CAASPP Coordinators, (b) school site CAASPP coordinators,[2] and (c) test administrators. The survey sample was designed to reflect the diversity of LEAs throughout the state, but *was not* intended to provide sufficient statistical representation of all LEAs to support precise statistical estimates of frequencies and proportions

[1] The California State Board of Education adopted the Common Core State Standards in August 2010.
[2] Hereafter, LEA CAASPP coordinators will be referred to as "LEA coordinators" and school site CAASPP coordinators will be referred to as "site coordinators."

for the state as a whole. These data mainly serve to illustrate the breadth and diversity of the experiences of the LEAs in the study. A full description of the development of data collection instruments, training of interviewers, response rates, demographics of respondents, survey and other study details and analyses can be found in Chapter 3.

What emerged from our study is a broad picture of a range of experiences from each of the three stakeholder groups surveyed. Though participation rates were disappointing and not generalizable statewide, the study offers an examination of local experiences of various factors related to the Smarter Balanced Interim Assessments that to this point has not yet been available. Overall, more than three-fourths of the respondents to each survey reported administering at least one Smarter Balanced Interim Assessment during the 2015–16 school year.

Major findings for each of the six research questions in the study, along with supporting evidence from results of three statewide online surveys and other statewide data, are presented here. Detailed descriptions are presented in Chapter 6.

Interim Assessment Study Research Question 1. How are decisions made about whether and how interim assessments are used?

The major finding provides an indication of how LEAs serving various populations use the interim assessments; however, due to our sample limitations, generalizations *cannot* be made as to how widespread this finding is.

> **Major Finding: In our sample, the primary factor that contributed to LEA and school decisions to require or highly encourage use of the Smarter Balanced Interim Assessments was the value of familiarizing students with the testing systems–including universal tools, designated supports, and accommodations–and the types of questions that are included in the Smarter Balanced Summative Assessments.** Nearly three fourths of LEA coordinators and of site coordinators who responded to our survey sample[3] and reported requiring or highly encouraging interim assessment usage cited these factors (78% of 236 and 75% of 87, respectively). The second most reported primary factor was to inform classroom instruction (59% of LEA coordinators and 64% of site coordinators).

Interim Assessment Study Research Question 2. What are detectable patterns in the types of interim assessments used (e.g., interim comprehensive assessments versus interim assessment blocks)?

> **Major Finding: Statewide, large numbers of schools and LEAs administered Smarter Balanced Interim Assessments during 2015–16.** ETS interim assessment usage data, through March 2016, show that 4,321,412 interim assessments were administered. Based on HumRRO's matching of interim assessment data to state-mandated summative assessment data, a total of 6,178 schools out of 9,212 (about 67%) indicated using at least one interim assessment, compared to only 3,034 schools that did not (33%).

[3] The survey sample of potential respondents included 1,599 LEA CAASPP coordinators; 1,006 school site CAASPP coordinators; and 12,751 test administrators. See Chapter 3 for sampling details.

Interim Assessment Study Research Question 3. To what degree were schools successfully prepared to administer the interim assessments (e.g., training materials were clear, system components worked)?

Major Finding: LEA coordinators in our sample generally found the various resources helpful and the various systems easy to use. Across all the training resources listed as response options, approximately half the LEA coordinators who responded to our survey found the resources to be helpful (50% of 285).[4] In particular, the caaspp.org resource links were perceived as helpful by nearly two thirds of LEA coordinators (63%). Across the various Smarter Balanced Interim Assessment systems, nearly half of LEA coordinators found the systems easy to use (46%). In particular, the Test Administrator Interface was the easiest to use (73%).

Interim Assessment Research Study Question 4. To what degree is the information about test administration procedures, as included in interim assessment resources, followed?

Major Finding: Most test administrators in our sample reported that their school attempted to standardize how the Smarter Balanced Interim Assessments were administered (e.g., required formal training on some components and required that procedures outlined in the *CAASPP Interim Assessment User Guide* be generally followed). Over four fifths of the test administrators who responded to our survey (83% of 863) reported that they followed the *CAASPP Interim Assessment User Guide* procedures to a moderate or extreme degree; a much smaller percentage (16%) indicated teachers were allowed great flexibility in administration.

Interim Assessment Study Research Question 5. To what degree do LEAs perceive the interim assessments impact instructional practice and student achievement?

Major Finding: LEA and school coordinators and test administrators who responded to our survey reported that Smarter Balanced Interim Assessments were useful in familiarizing students and teachers with the format and content of the Smarter Balanced Summative Assessments, which are mandated by the state. Familiarizing students and teachers with the testing system, item types, and scoring rubrics was the most useful benefit reported by all three stakeholder groups (average for LEA coordinators = 72% of 285, average for site coordinators = 73% of 90, average for test administrators = 54% of 863). This finding is consistent with analysis of statewide assessment data, which shows that schools that administered the Smarter Balanced Interim Assessments (either blocks or comprehensives) had more improvement in school-level summative scale scores than those that did not administer the interim assessments.

[4] For Likert scale items presented in a matrix format with several rows of response options, we provide the sample's average for the set of options.

Interim Assessment Study Research Question 6. What challenges existed in the 2015–16 school year that could be improved for 2016–17?

> **Major Finding: LEA and school site coordinators and test administrators in our sample did not report significant challenges with Smarter Balanced Interim Assessments.** On average, about four fifths of LEA coordinators who responded to our survey rated various administration activities (e.g., bandwidth delays) as not challenging or only a minor challenge (81% of 863). On average, about three fifths of LEA coordinators rated test content features (e.g., difficulty of content) as not challenging or only a minor challenge (61%). On average, nearly half of LEA coordinators rated assessment reporting activities as not challenging or a minor challenge (46%).

Based on the major findings from the Smarter Balanced Interim Assessment Administration Study, we offer the following recommendation:

> **Interim Assessment Study Recommendation: Continue to monitor the various Smarter Balanced Interim Assessment systems and components. We recommend that the CDE also continue collecting feedback from schools and LEAs, as well as from other Smarter Balanced Assessment Consortium members, to see where reasonable improvements could be made to the system.**

Major Findings and Recommendation from the Access to Designated Supports and Accommodations Study

We conducted the Access to Designated Supports and Accommodation Study to (a) examine the availability and use of testing supports and accommodations for students with disabilities (SWDs) and English learners (ELs) on the Smarter Balanced ELA and mathematics assessments and (b) determine whether the tools used for these assessments are consistent with those used routinely by students in their classrooms. Our study obtained a sample of LEAs, schools, and local staff that was about one-half of the originally targeted sample size (6 of 12 LEAs and 15 of 36 schools). The reduced sample sizes allowed us to collect information from the study participants in greater depth, leading to what was more like a series of case studies. Although there are significant limitations with respect to the generalizability of this study's findings, results from the study may inform future larger-scale studies on this topic.

We generated qualitative and quantitative data to address the study's research questions. The qualitative data were gathered in phone-based focus groups, in-person interviews with local staff during school site visits, and observations of instruction and assessment. Quantitative data related to the use of instructional accommodations were captured from teacher reflection checklists and from items on the survey conducted as part of the Interim Assessment Administration Study. We also analyzed extant data from the 2016 Smarter Balanced Summative Assessment, relative to accessibility features offered to students. A full description of the development of data collection instruments, training of interviewers, study participants, and other study details and analyses can be found in Chapter 4.

Though the sample is very small, it includes a range of LEA and school characteristics. Anecdotal evidence indicated that some LEAs were hesitant to participate if they felt their staff lacked familiarity with issues related to testing accommodations. Participation within LEAs also proved to be a challenge. For example, only about half of teachers who agreed to complete reflection checklists actually did so.

Major findings for each of the following four research questions in the study, along with supporting evidence, are based on the limited number of LEAs, schools, and teachers who participated in the study. Findings based on such small sample sizes should be considered just that – reflective of the responses in the small sample – and not be interpreted to be representative of the state-level population of LEAs. Expanded descriptions of the findings are presented in Chapter 6.

Access to Designated Supports and Accommodations Study Research Question 1. Is the general assessment accessible to moderately disabled students and English learners through the provision of accommodations and supports?

The Smarter Balanced Summative Assessments include a wide variety of universal tools, designated supports, and accommodations designed to allow students to access the assessments and demonstrate what they know and can do. Within the new online testing environment, the entire suite of sophisticated designated supports and accommodations that students can be offered is a tremendous expansion beyond what has been available in paper and pencil mode.

The online CAASPP Smarter Balanced assessments allow universal tools, designated supports, and accommodations (embedded and non-embedded). ***Universal tools*** are available for all students based on student preference and selection, and include, but are not limited to, breaks and digital notepad. ***Designated supports*** are available to *all* students when determined for use by an educator or group of educators (with parents/guardian and student input, as appropriate) or specified in the student's individualized education program (IEP) or Section 504 plan. These include but are not limited to color contrast and separate setting. ***Accommodations*** are available if specified in the student's IEP or Section 504 plan and, among other resources, include braille and American Sign Language for Writing, Listening, and Mathematics.[5]

> **Major Finding: Students with disabilities and English learners were offered a wide range of accessibility features during interim and summative assessments, although actual use of these features cannot be determined.** Population-level data on summative assessment accommodations indicate that all students with disabilities and English learners were provided access to at least one accessibility feature through the provision of accommodations, designated supports, and universal tools. The assessment software does not capture student use of features, however.

Access to Designated Supports and Accommodations Study Research Question 2. To what extent do the supports and accommodations provided and used in the interim and summative assessments match those used in classroom instruction for individual students?

Using a checklist developed by HumRRO and based on the *Smarter Balanced Resources and Practices Comparison Crosswalk*, teachers in our study sample documented their instructional practices and classroom assessment (i.e., formative, diagnostic, interim, or benchmark testing) for individual students. They also recorded accommodations they used with these students over the course of the 2015–16 school year. HumRRO then merged teacher checklist data with data on summative assessment accommodations and designated supports and created indicators of when students received both an instructional practice and the corresponding summative assessment accessibility feature. These data are supplemented with focus group and site visit data, including eight observations of summative assessment administrations.

[5] Source: http://www.cde.ca.gov/ta/tg/ai/caasppmatrix1.asp (Retrieved 10/5/16).

Major Finding: In our sample, assessment accessibility features offered to students were generally aligned with instructional practices of the students' teachers. Among the most frequently used instructional practices indicated in the small matched sample of students, the majority of students typically were offered during assessment the associated accessibility feature, most often through the provision of universal tools.

Access to Designated Supports and Accommodations Study Research Question 3. Are there types of supports or accommodations used by students when learning in the classroom that are not used on assessments?

Major Finding: In our sample, teachers rarely reported using instructional practices that were not reflected in the summative assessment accommodations, designated supports, and universal tools available. When given the opportunity to identify "Other" instructional practices, teachers in our sample who completed checklists tended to identify practices that were already listed, or that could not be provided during testing without changing the construct.

Access to Designated Supports and Accommodations Study Research Question 4. How often do students attempt to use test supports and accommodations that they do not use in classroom instruction?

Major Finding: There does not appear to be a widespread issue of students being offered large numbers of unneeded accommodations and designated supports on the summative assessments. Population-level data show that students with disabilities and English learners were typically offered one to two accommodations or designated supports. Very few students were given large numbers of accommodations or supports. The sample of educators and testing coordinators who participated in our focus groups tended to express wariness of offering accommodations or designated supports that were inappropriate and thereby might invalidate students' scores.

Based on the major findings from the Access to Designated Supports and Accommodations Study, we offer the following recommendation:

Access Study Recommendation: Continue to monitor the accessibility features available in the Smarter Balanced assessment systems to see where reasonable improvements could be made. Continue to collect feedback from stakeholders to identify and promote best practices for implementing accessibility features.

Major Finding and Recommendation from Observations of CAASPP Training Sessions

Based on our observations of seven training sessions conducted during 2015–16, we summarize findings and supporting evidence regarding professional development resources for the CAASPP System.

Major Finding: The CAASPP Institutes and Post-test Workshop provide high quality in-person training and valuable reference materials for remote access by teachers and other educators. Presenters communicated substantive content about targeted aspects of the CAASPP System to a variety of types of educators with differing perspectives on the CAASPP System and differing levels of training and experience in

assessment theory and implementation. Materials included a variety of graphics and other content organizers to illustrate and emphasize essential concepts. The CDE posted links to electronic files of all materials from these sessions.

Based on the findings from the observations of professional training, we offer the following recommendation:

Training Sessions Recommendation: Continue to support professional development opportunities and maintain online resources that enhance LEA and school staff understanding of how best to utilize all components of the CAASPP System to improve teaching and learning.

General Summary

Our study of interim assessments found wide usage of the CAASPP interim assessments. Initial use of these assessments focuses on familiarizing students with the summative test and testing environment. The use of interim assessment results to improve instruction will likely increase over time. Current evaluation work has resulted in a number of suggestions for improving the efficiency and efficacy of the use of interim assessments.

Our study of supports and accommodations found good alignment of supports and accommodations during testing for students with disabilities and English Learners with those used routinely in classroom instruction. A chief concern was that some students may not understand how to respond to some of the newer item types. Given this concern, the use of the interim (and practice) tests to familiarize students with the testing environment is entirely appropriate.

Overall, results to date indicate the California public school system has implemented the Smarter Balanced assessments in a way that provides stable and meaningful results.

This page is intentionally blank.

CHAPTER 1: INTRODUCTION

Background

The California Assessment of Student Performance and Progress (CAASPP), established by California Assembly Bill (AB) 484, became the state's new student testing program on January 2, 2014. The CAASPP replaced the former Standardized Testing and Reporting (STAR) Program that California had in place since 1998. This new testing program shifts the focus for assessments away from accountability towards a comprehensive plan for promoting high-quality teaching and learning for all students, including students with disabilities (SWDs) and English learners (ELs).

AB 484 requires an independent evaluation to provide objective technical advice and consultation on activities to be undertaken in implementing the CAASPP System. The evaluation is defined in California *Education Code* (*EC*) Section 60649, which states that evaluation "activities may include, but not necessarily be limited to, 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 California Department of Education (CDE) awarded the Human Resources Research Organization (HumRRO) a three-year contract to conduct an independent evaluation of the CAASPP System beginning July 2015 and concluding December 2017. At the CDE's direction, the scope of the current independent evaluation is limited to the following new assessments that are being implemented as part of the CAASPP System:

- Smarter Balanced Summative Assessments for English language arts/literacy (ELA) and mathematics, administered in grades three through eight and high school
- Smarter Balanced Interim Assessments for ELA and mathematics, designed for grades three through eight and high school, available to educators who teach grades kindergarten through grade twelve, optional for use by local educational agencies (LEAs) and teachers at non-public schools who provide direct instruction to California public school students
- California Alternate Assessments[6] for ELA and mathematics, given in grades three through eight and grade eleven
- California science assessments[7] including alternate assessments, in grades to be determined
- Primary Language Assessments for Reading/Language Arts (RLA)[8] in grades three through eleven, optional for use by LEAs

[6] California Alternate Assessments are successors to California Alternate Performance Assessment (CAPA), ELA and mathematics.
[7] Science assessments - the California Science Test and the California Alternate Assessment for Science are the successors to the California Standards Test (CST), California Modified Assessment (CMA), and CAPA for Science.
[8] Primary Language Assessments for RLA are successors to Standards-based Tests in Spanish (STS).

The evaluation contract calls for three annual reports of independent evaluation activities, findings, and recommendations. Within a few months of contract award, HumRRO submitted to the CDE the first required annual evaluation report (Hardoin, Becker, Wise, 2015). That report's core contents included the three-year evaluation plan approved by the State Board of Education (SBE) and an overview of the five research studies to be conducted during the contract period. The present report is the second annual report, and it describes results from the 2015–16 studies. The third annual report will describe the results from the 2016–17 studies. A Comprehensive Final Evaluation Report will be delivered in 2017 and will include evaluation findings from each of the three annual reports (2015, 2016, and 2017). We provide a brief summary below of the activities reported in 2015 as context for the continuing evaluation activities.

Summary of 2015 Evaluation Activities

Initial activities related to the independent evaluation involved development of a three-year evaluation plan in cooperation with CDE staff, the Technical Advisory Group (TAG), and SBE members and staff. The work began with a theory of action that, along with CDE priorities and the timeline for administration of the assessments, guided the selection and sequence of five research studies to be conducted over the course of the contract period.

The design of the evaluation research studies began with the premise of the CDE's overall goal for implementing the CAASPP System of formative,[9] interim, and summative assessments, which is to improve the effectiveness of instruction and student effort and engagement, and thereby increase student achievement. In consultation with the CDE, HumRRO established research studies to collect evaluative evidence that would support or identify areas of strength or weakness related to the following assumptions essential to the system's theory of action:

1. Test results provide accurate and unbiased assessments of what students have and have not learned. An accurate assessment is valid, fair, and reliable for all groups of students. The *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 2014) assert that the validity of score interpretations for their intended use is primary. Issues of fairness and reliability support the validity of interpretations, and validity concerns are important in each step in the testing process from initial design through reporting. Evaluation of these components of the system involves evaluation of test content, access issues, assessment administration, scoring, and reporting of test results.

2. Test result information is used in ways that improve the effectiveness of instruction, including use by (a) teachers to monitor and improve their practices and to target instruction for individual students, (b) students and parents to seek remedial help where needed and to increase motivation for students who are behind, and (c) policy makers of school accountability information based on test results to focus school improvement efforts and/or to initiate restructuring efforts as needed. The *Standards for Educational and Psychological Testing* also describe the need to investigate the impact or consequences of test use. It is important to test assumptions in the theory of action that provide the rationale for developing and using the test for specific purposes.

[9] Formative assessment processes and tools, including the Smarter Balanced Digital Library and other available resources for teachers, are outside the scope of the current independent evaluation.

We presented five studies for consideration at the summer 2015 TAG meeting. For each study, we provided the rationale for the research, the research questions to be answered, and an overview of the methods to be used and data to be collected. Although the studies were mostly targeted to evaluating whether test results are accurate and valid for their interpretations and use, the studies also included exploration of the extent to which test result information is used to monitor and improve instructional practice. Discussion of the evaluation studies included participation of TAG members, CDE staff from the Assessment Development and Administration Division (ADAD) and the Accountability, Measurement, and Reporting Division (AMARD); Educational Testing Service (ETS); and the SBE. Meeting attendees asked questions about and provided comments on the proposed studies; their input was used to narrow and prioritize the focus of HumRRO's studies and develop the three-year evaluation plan. The planned independent evaluation studies, in accordance with the law, avoided duplication of studies already planned as part of federal peer review or by assessment contractors.

During the September 2, 2015 SBE meeting, CDE Deputy Superintendent Keric Ashley reviewed legislation that requires the independent evaluation of the CAASPP System and presented HumRRO's CAASPP Independent Evaluation Study Plan. The SBE approved the plan, recommending that the plan be reviewed and updated as needed to address new high-priority issues. Outcomes related to the 2016 review of the original CAASPP Independent Evaluation Study Plan are presented in a later section of this report.

HumRRO's CAASPP 2015 Independent Evaluation Report included the SBE-approved CAASPP Independent Evaluation Study Plan in its entirety. The three-year plan described rationales and research questions for the following five independent evaluation studies:

- Interim Assessment Test Administration Study, to be conducted during the 2015–16 school year
- Access to Designated Supports and Accommodations Study, to be conducted during the 2015–16 school year
- Review of Scoring Processes Study, to be conducted during the 2016–17 school year
- Use of Assessment Results Study, to be conducted during the 2016–17 school year
- Item Alignment and Quality of Science Assessments Study, to be conducted during the 2017–18 school year

In addition to developing and receiving approval of the three-year evaluation plan, we also established connections with LEAs. As the independent evaluator, we anticipated that data collection activities associated with the evaluation studies would include focus groups, interviews, and classroom observations with California educators. To facilitate support from and participation of LEA and school staff in our research, we created a collaborative relationship with a sample of LEAs (the Local Educational Agency Research Network or LEARN). The sample, selected in consultation with the CDE, was designed to represent a variety of LEAs by size (in terms of student population served), geographic region, and other characteristics. This group of LEAs is intended to provide an "on the ground" perspective regarding the CAASPP System as well as participate in specific studies. We plan to maintain our relationship with LEARN members throughout the course of the three-year evaluation and replace or supplement membership as needed.

Although the CAASPP independent research studies were not yet underway when the 2015 Evaluation Report was produced, we offered the following recommendation about the LEARN in the first annual evaluation report:

Recommendation 1: Establish the Local Educational Agency Research Network (LEARN) as soon as possible. The guidance and participation of a representative sample of LEAs working in collaboration with HumRRO will be essential to the two studies selected for the first evaluation year: Access to Designated Supports and Accommodations, and Interim Assessment Test Administration. The recruiting of LEARN members is designed as a joint effort between the CDE and HumRRO, in that the CDE will issue a letter to superintendents and CAASPP coordinators of selected LEAs to encourage participation, and HumRRO will recruit LEAs to engage them in the evaluation research. This two-step process will be repeated as needed to replace any selected LEAs that decline to participate. The sooner this network is established the sooner study activities can proceed.

Organization and Contents of the CAASPP 2016 Evaluation Report

This 2016 Evaluation Report covers activities performed from the date of SBE approval of the CAASPP Independent Evaluation Study Plan through completion of the first two research studies, the Interim Assessment Test Administration Study and the Access to Designated Supports and Accommodations Study. The report includes summaries of additional project start-up activities; detailed designs for both studies, data collection activities for both studies, and findings from both studies; as well as outcomes of the 2016 review of the evaluation plan.

Chapter 2 of this report, *Implementing the Independent Evaluation Plan*, presents key activities we conducted to launch the SBE-approved CAASPP Independent Evaluation Study Plan. Activities include the creation of a collaborative research network (the LEARN) with a targeted sample of 24 LEAs, background research on the CAASPP System, and in-person observations of CAASPP System training provided to California teachers, administrators, and school district staff by the assessment contractor or by the Sacramento County Office of Education on behalf of the CDE. Additionally, we describe policies and procedures implemented to ensure the security and confidentiality of CAASPP evaluation data.

Chapter 3, *Smarter Balanced Interim Assessments Administration Study*, presents the research questions, methods, and data collection activities conducted to investigate California LEA and school educator experiences with the Smarter Balanced Interim Assessments during the 2015–16 school year. Because interim assessments are administered at the discretion of LEAs rather than mandated by the state, we designed the study to gain greater understanding of how this component of the CAASPP System works. The multi-staged study began with a small number of focus groups and interviews to explore how and why interim assessments were used; the study concluded with administration of three statewide online surveys. Our online interim assessments administration surveys were tailored to three different roles: LEA CAASPP Coordinator, Site Coordinator, and Test Administrator. The very low response rates to the surveys, particularly for the test administrators, preclude us from interpreting findings as generalizable statewide; however, the responses from approximately 1,300 educators do suggest variability in policies for using Smarter Balanced Interim Assessments, as well as a range of educator perceptions of preparation and training for administering the assessments, impacts of the assessments on instruction, and areas for improvement in future years. The chapter presents analysis of focus group data and summary findings that integrate the qualitative and quantitative data analysis of the survey data, organized by major topics.

Chapter 4, *Access to Designated Supports and Accommodations Study*, presents the research questions, methods, and data collection activities conducted to investigate the availability and use of tools, designated supports, and accommodations during classroom instruction and when

students complete assessments. The goals of the study were to determine whether there is consistency between the two settings (i.e., instruction and assessment) and to identify any inconsistencies that could be potential threats to the validity of test score interpretation and use for SWDs and ELs. For this study, we conducted a small number of focus groups and interviews with LEA staff and teachers to explore the relationship between accommodations and supports used in daily instruction and those used on assessments, including large-scale assessments. During school site visits, HumRRO researchers observed firsthand the use of instructional and assessment supports and accommodations. Site visits also were used to train educators to complete a self-reflection checklist of 2015–16 instructional and assessment practices with SWDs and ELs. Finally, we analyzed the 2016 CAASPP Smarter Balanced summative assessment data to further investigate the accommodations and designated supports offered to students statewide and to students for whom we had checklist data. Although the number of LEAs and schools that participated in this study was quite small and the sample was not representative of the state as a whole, the findings illustrate variability in teacher understanding and training in the use of tools, designated supports, and accommodations on summative assessments. The chapter presents summary findings of qualitative and quantitative data analysis.

Chapter 5, *Update to the CAASPP Independent Evaluation Study Plan*, describes the review and revision of the original CAASPP Independent Evaluation Study Plan. The review was conducted to respond to contextual issues that were not fully anticipated in the original design, namely the changing timelines of the development and administration of the California Alternate Assessments and the Science Assessments. The chapter presents a summary of the three remaining studies to be conducted under the Updated Independent Evaluation Study Plan (2016), including an alignment study of the California Alternate Assessments.

Finally, Chapter 6 presents our findings and recommendations based on the data analyses and results presented in preceding chapters.

This page is intentionally blank.

CHAPTER 2: IMPLEMENTING THE INDEPENDENT EVALUATION PLAN

Michele Hardoin, Sunny Becker, and Monica Gribben

HumRRO began implementing its three-year plan for conducting independent research into components of the California Assessment of Student Performance and Progress (CAASPP) System shortly after approval of the Independent Evaluation Study Plan by the California State Board of Education (SBE) on September 2, 2015. This chapter presents four critical activities we completed prior to or in concurrence with the two specific research studies we conducted during the 2015–16 school year. Here, we report on (a) background research into the CAASPP Smarter Balanced assessments, (b) our commitment to security and confidentiality for all evaluation data handled by HumRRO's CAASPP project team, (c) observations of training sessions conducted by the assessment contractor or by others to orient California educators to various aspects of the CAASPP system, and (d) the establishment of a collaborative research network with a representative sample of local educational agencies (LEAs) that will support the ongoing work of the evaluation. The 2015–16 research studies specifically address the Smarter Balanced English language arts/literacy (ELA) and mathematics assessments in the CAASPP System. We present these studies in detail as follows: in Chapter 3 we report on *Smarter Balanced Interim Assessments Administration Study,* and in Chapter 4, *Access to Designated Supports and Accommodations Study*.

Background Research on CAASPP System

The 2014–15 CAASPP System included the Smarter Balanced assessments and the Smarter Balanced Digital Library. To build our knowledge and understanding of the CAASPP System and to understand how the system's components are presented to California teachers, administrators, and district staff, we reviewed information available to the public online as well as resources made available by the California Department of Education's (CDE) technical monitor. This foundational knowledge was essential to our development of appropriate data collection instruments for our research studies.

The CDE arranged for HumRRO's key project staff to obtain Digital Library user accounts, from which we obtained a wealth of information about specific aspects of assessment administration and accessibility.[10] Project staff also obtained information through subscriptions to CDE's *CAASPP Update*, a weekly e-mail that provides highlights of the CAASPP System, including the Smarter Balanced assessments, the California Alternate Assessments, and science assessments. We studied 2014–15 and 2015–16 CAASPP informational videos of Webcasts and supporting materials, test administration manuals, tools (e.g., the Individual Student Assessment Accessibility Profile ISAAP tool), the accessibility matrix, and other program documents. We also examined revised versions of existing resources as they became available.

Safeguarding Confidential Data

HumRRO fully understands the importance of adhering to policies that protect and monitor access to sensitive information, such as student level data, while carrying out our independent evaluation. We 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.

[10] We obtained this information by accessing the CAASPP portal at caaspp.org.

For the CAASPP evaluation, our security program focuses 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 security training on specific security requirements related to our 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 security procedures are administered in an efficient and effective manner. We describe the details of our staff training and procedures implemented to address these key areas in Appendix A1.

Observation of CAASPP Educator Training Sessions

To supplement the online resources and to gain firsthand knowledge of training that LEA and school staff received about the CAASPP Smarter Balanced assessments—in particular those areas related to our 2015–16 independent studies—HumRRO staff observed a total of seven training sessions throughout the state over the course of the school year.[11] Each type of training session we attended was offered multiple times and conducted at different locations in the state; additional types of training sessions were offered that were not observed by HumRRO staff. We attended all sessions as observers, and we informed session facilitators that we represented the independent evaluator. We 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, our key findings, and several recommendations for improving the effectiveness of the training.

The following five sessions were designed to be professional development opportunities for public school educators:

- *Fall 2015 CAASPP Institute (Part 1 of a two-part training) Train-the-Trainer,* conducted by the Sacramento County Office of Education (SCOE) in partnership with CDE. To support LEAs and schools that were unable to attend the initial offering of the CAASPP Institutes, this one-day session provided the knowledge, training, and materials for County Offices of Education to deliver comprehensive CAASPP training (i.e., training addressing the Smarter Balanced Digital Library, Summative, and Interim Assessments) within their counties.
- *Spring 2016 CAASPP Institute (Part 2 of a two-part training),* conducted by the Sacramento County Office of Education in partnership with CDE. This was a one-day regional meeting for multi-disciplinary, three- to six- member teams of LEA and school educators. This meeting served as the second part of a professional learning opportunity during which teams were taught strategies for using the CAASPP System to improve teaching and learning. Participants were provided the opportunity to share ideas and challenges and collaborate with colleagues.
- *2015–16 Post-test Workshop: Connecting Assessments to Instruction,* conducted by Educational Testing Service (ETS), WestEd, and the CDE Assessment Fellows. The session provided information on using the Online Reporting System and available reports, including Assessment Target Reports, to inform and improve classroom instruction. LEAs were encouraged to bring LEA CAASPP coordinators, professional development staff, and curriculum specialists.

[11] One or more HumRRO staff members participated in each event at one of the locations where they were offered.

- *Fall 2015 Digital Library and Interim Assessment Clinic,* conducted by staff from ETS. This session taught educators the procedures for accessing the Digital Library and provided information about the resources designed to support teaching and learning as well as which Digital Library resources support the formative assessment process. Information about the Smarter Balanced Interim Assessments was also provided, including how to administer them and considerations for using the tests and results to enrich and inform teaching.
- *Fall 2015 Interim Assessment Hand Scoring Workshop,* conducted by ETS. This session informed educators about the general interim assessment administration process and introduced the concepts and processes for scoring open-ended student responses. Educators were trained in the effective use of the hand scoring training guides and exemplars to accurately evaluate student work.

The following two sessions HumRRO staff attended were in-person versions of training that was also delivered online to certify California public school educators and other content specialists as hand scorers for the 2016 Smarter Balanced assessments:

- *Spring 2016 Smarter Balanced Summative Assessment Hand Scoring Training,* conducted by Measurement Incorporated (MI). Educators received hands-on training and had the opportunity to become certified raters. The sessions were organized by content area and grade span; HumRRO attended the training for ELA grades six through eight.
- *Spring 2016 Smarter Balanced Summative Assessment Hand Scoring Training,* conducted by ETS. Educators received hands-on training and had the opportunity to become certified raters. The session was organized by content area and grade span; HumRRO attended the training for mathematics grade eleven.

Creation of a Collaborative Research Network with Local Educational Agencies

As the independent evaluator for the California High School Exit Examination (CAHSEE), we collaborated with research and assessment staff from LEAs in California on a research study that explored the relationship between post-high school outcomes and students' performance on the CAHSEE (Becker, Wise, Hardoin, & Watters, 2012). Based on this successful collaborative experience, we devised a similar approach to conduct the CAASPP evaluation.

Some unique contributions proposed for the CAASPP evaluation include exploring aspects of the Smarter Balanced Interim Assessments. Unlike the summative assessments mandated by the state, the determination of which, if any, interim assessments to use is left to the discretion of the LEAs. Participation and support of LEA research staff was essential when conducting our research on interim assessments. In addition, data collection for other evaluation studies, such as the 2015–16 Access to Designated Supports and Accommodations Study, involve focus groups and classroom observations, activities that require assistance from LEA staff who could connect us to educators in the LEA's schools. We envisioned a network of LEAs that represent the diversity of the state and works with us to support the evaluation in ways that will be mutually beneficial to the participating LEAs, HumRRO, and the CDE. For the duration of the evaluation, we will refer to this representative group of member LEAs as the Local Educational Agency Research Network, or the LEARN.

Selection and Recruitment

Recruiting LEARN members was designed as a joint effort between the CDE and HumRRO. To begin the process, HumRRO obtained from CDE staff in August 2015 a statewide database with contact information for all CAASPP LEA coordinators. We combined this information with 2015 CAASPP Smarter Balanced ELA and mathematics results. This combined file served as the basis for characterizing schools for possible selection. These data were used to empirically derive cut points to classify each LEA along four dimensions: (a) small, medium, large, or very large size in terms of student enrollment; (b) high or low percentage of ELs; (c) high or low academic performance as determined by a HumRRO-created academic index; and (d) north or south geographic region based on LEA zip code.

1. **Size classification** based on number of students with 2015 CAASPP records: Small = < 1,000 students, Medium = 1,000 to 8,999 students, Large = 9,000 to 34,999 students, Very Large = 35,000 or more students. Our sampling frame did not include student counts by County Office of Education (COE) so all COEs were coded as "Missing."
2. **Percent of EL students classification**: Low = 0 to 15 percent EL, High = more than 15 percent ELs.
3. **Academic index:** HumRRO calculated a summary statistic to categorize LEAs as high or low performing.[12]
4. **Geographic region**: South = Zip code < 93000, North = Zip code ≥93000

We selected LEAs using these four criteria. Per CDE request, we selected two charter schools and two COEs; due to unavailability of some sampling frame data these were selected solely on the basis of geographic region. The remaining LEAs were selected on the basis of all four criteria. We selected 24 primary LEAs and 24 alternate LEAs, whereby each primary LEA had an alternate LEA that mirrored the characteristics of its counterpart and would serve as a substitute if the primary LEA declined participation in the LEARN. HumRRO submitted the two lists of 24 LEAs, including all four very large LEAs, two COEs, and two direct-funded charter schools, to the CDE on September 23, 2015. The CDE approved the primary and alternate lists of LEAs for recruitment.

To begin the communication process with selected LEAs, the Director of the Assessment Development and Administration Division at the CDE issued an e-mail to the superintendents and charter school administrators from the primary list on October 13, 2015, with a copy to CAASPP coordinators. This e-mail briefly explained the LEARN and endorsed LEA participation (see Appendix B1). HumRRO sent a follow-up e-mail to the CAASPP coordinators on October 21 that included additional detail about the expected effort required to be a LEARN member, a commitment form, and a copy of the SBE-approved evaluation plan (see Appendix B2).

When an LEA declined to participate, we notified the CDE to send the initial e-mail to the corresponding alternate LEA, after which we followed up with that alternate LEA to gain its participation. This two-step process was repeated as needed to replace any LEAs that declined

[12] We converted grade-level mean scores for each test to z-scores (school Avg. – State Avg.)/State standard deviation [SD]. We then computed a weighted average of the z-scores across grades and subjects, weighted by the number of students with data for each grade and subject. The academic index values were: High = Academic z-score > 0, Low = Academic z-score 0 or less, and "missing" if we had insufficient numbers to compute an average. LEAs with a missing academic index were excluded from sampling.

to participate. We successfully secured 24 LEA members, 13 from the primary list and 11 from the alternate list, reaching the target number of recruits as of November 18, 2015.

Membership

Each LEA that joined the LEARN designated a primary point of contact for HumRRO to work with on the independent evaluation activities. By joining the LEARN, LEAs committed to actively participate in a half-day Web-based meeting approximately one or two times per year. In these meetings HumRRO reviewed our research plans and preliminary results, and LEARN representatives commented on the feasibility of plans and helped interpret findings. In addition to these meetings, LEAs agreed to participate in one or more research studies over the course of the evaluation. Participating in studies was described as potentially responding to surveys, participating in interviews or focus groups, allowing observations of test administrations and scoring, providing assessment or other data, or other activities. During the recruitment process we explained that details would be provided as study plans matured.

The following 24 LEAs committed to joining the LEARN:

Alum Rock Union Elementary
Clovis Unified School District
Corona-Norco Unified School District
Elk Grove Unified School District
Fairfield-Suisun Unified School District
Fontana Unified School District
Fresno County Office of Education
Fresno Unified School District
Learning for Life Charter
Long Beach Unified School District
Los Angeles Unified School District
Mammoth Unified School District
Merced City Elementary
Mt. Diablo Unified School District
Oakland Unified School District
Poway Unified School District
San Diego County Office of Education
San Diego Unified School District
Santa Monica-Malibu Unified School District
Tulare City
Valle Lindo Elementary
Victor Valley Union High
Westchester Secondary Charter
Westside Union Elementary

HumRRO wishes to thank the LEARN member representatives and the teachers and administrators at their schools who participated in the LEARN. We especially thank those who participated in the Smarter Balanced Interim Assessment Study and the Access to Designated Supports and Accommodations Study. Our research would not have been possible without their commitment and involvement.

Table 2.1 presents the classification of the LEARN members per the selection criteria.

Table 2.1. Classification of 2015–16 LEARN Members by Selection Criteria

LEA	Size	Percent EL	Region	Academic Performance
1	Small	High	North	High
2	Small	Low	South	High
3	Small	n/a	North	n/a
4	Small	n/a	South	n/a
5	Medium	High	North	High
6	Medium	High	North	Low
7	Medium	High	North	High
8	Medium	High	North	Low
9	Medium	High	South	Low
10	Medium	Low	North	High
11	Medium	Low	South	Low
12	Medium	Low	South	High
13	Large	High	North	High
14	Large	High	North	High
15	Large	High	South	High
16	Large	Low	North	High
17	Large	Low	North	Low
18	Large	Low	North	High
19	Large	Low	South	Low
20	Large	Low	South	Low
21	Very Large	High	North	Low
22	Very Large	High	South	High
23	Very Large	High	South	High
24	Very Large	High	South	Low

Note. Size is categorized as Small < 1,000 students; Medium = 1,000 – 8,999 students; Large = 9,000 – 34,999 students; Very Large = 35,000 or more students. *Percent EL* is defined as Low = 0 – 15% EL students; High = 15% or more EL students. *Geographic region* is classified as South = zip code < 93000; North = zip code of 93000 or greater. *Academic Performance* is based on a HumRRO-generated variable consisting of a weighted average of z-scores across grades and subjects on 2015 Smarter Balanced ELA and mathematics assessments. Academic Performance is categorized as Low if less than or equal to zero and High if greater than zero.

Inaugural Meeting

When LEARN membership was finalized, the school year was nearly half complete. HumRRO aimed to launch the LEARN activities as soon as possible to engage LEAs in the current year's studies. We polled LEARN representatives for their availability to participate in a three-hour Web-based meeting to introduce them to the overall CAASPP independent evaluation and the two research studies being conducted in 2015–16. The first inaugural LEARN meeting was conducted on December 4, 2015, with a second conducted on December 8 for those unable to attend the first meeting. In total, 19 LEA representatives attended one of the two inaugural meetings. Participants received a meeting agenda and presentation materials as well as fliers that described each of the 2015–16 research studies, including estimates of the time and effort expected of LEA and school staff participants.

During each three-hour meeting, CDE staff welcomed and thanked the LEARN representatives for their time and commitment to the evaluation effort. HumRRO facilitated introductions of the members, provided an overview of the CAASPP evaluation, described expectations of the LEARN, and presented detailed descriptions of the two studies underway in 2015–16. Members participated actively during both sessions. We recorded the second Webinar and provided the link and electronic file of the presentation to all LEARN representatives, enabling those who were unable to attend either meeting to experience the orientation. We e-mailed the LEARN membership list, meeting feedback form, and study commitment form to each LEARN representative with a request to return completed forms by December 31, 2015.

Further details about LEARN member recruitment and participation in the 2015–16 research studies are provided in later chapters that focus on each study.

Spring 2016 Meeting

HumRRO LEARN representatives were polled on February 26, 2016 to select a date for our second meeting; reminders were sent to non-respondents on March 4 and 18. We selected the date that afforded the highest possible meeting attendance.

The spring 2016 meeting allowed us the opportunity to review initial findings of the 2015–16 studies and introduce and gain preliminary LEARN member participation in the 2016–17 research studies. We e-mailed an agenda and meeting slides to LEARN member representatives and CDE staff prior to the meeting. We conducted the Web-based meeting with nine LEARN representatives, one CDE staff member, and five HumRRO staff on May 13, 2016. Attendees participated via phone while watching a PowerPoint presentation online. Participants could enter comments or questions into a Web-based comment window, as well as provide them verbally.

HumRRO study leaders described progress on the two 2015–16 studies: the Access to Designated Supports and Accommodations Study and the Interim Assessments Administration Study. They also presented HumRRO's initial findings, based on data and analyses to date. LEARN participants engaged in an active discussion of these findings, offering alternative explanations for some observations and suggesting additional issues to be considered. These discussions are not detailed here, as they are subsumed within the study results in chapters 3 and 4 of this report.

The discussion then moved to upcoming studies. HumRRO study leaders presented plans for the 2016–17 studies: the Review of Scoring Process Study and the Utility of Score Reporting Study. Presenters described preliminary study plans, subject to review and approval by the CDE. We asked LEARN member representatives for feedback on the study designs, including whether some aspects of the design would be feasible or could be improved. LEARN representatives provided useful suggestions that subsequently were incorporated into the study plans. We also asked for a nonbinding indication of whether these LEAs would be willing to participate in either or both of these studies, and responses were quite encouraging.

Our final agenda topic was to discuss the effectiveness of the LEARN. We noted that, while attendance at the spring 2016 meeting was disappointingly low, the participating LEARN representatives were actively engaged and HumRRO found the feedback they provided to be very useful. HumRRO pointed out the reduced engagement of some LEARN members including low response rates to polls for potential meeting dates as well as difficulty recruiting LEAs to participate in the 2015–16 studies.

The group engaged in a healthy and frank discussion, and LEARN members postulated several reasons for low study participation. These include concerns that staff felt their lack of knowledge about aspects of the assessments would be highlighted, confusion about the differences between summative and interim assessments, and concern that accommodations and supports were not being provided appropriately. In addition, potential participants felt unable to commit the perceived amount of time required for these studies.

Recommendations for Maintaining an Active LEARN

We make the following recommendations for our continuing work with the LEARN, incorporating suggestions made by LEARN member representatives during the spring 2016 meeting:

- *Recruit for studies earlier in the school year.* Due to the time required to recruit and assemble the LEARN, recruitment in 2015–16 did not begin until December 2015. Recruiting in late summer–early fall might yield better results.
- *Provide more effective recruitment fliers:* HumRRO provided brief (one sheet, double-sided) documents for each 2015–16 study for LEARN representatives to use to obtain LEA agreement to participate and to recruit district and school staff. In the interest of completeness, these fliers were dense with information, including the expected amount of effort required to participate in each stage. LEARN representatives explained that this format was off-putting and suggested we provide a more user-friendly flier, along with a separate document with all the details. Several LEARN representatives offered to review preliminary versions of these fliers and provide feedback.
- *Increase ongoing communication with the LEARN:* Participants agreed that twice-yearly meetings were not too heavy a burden; some noted that quarterly meetings would be acceptable. Attendees suggested that periodic e-mail updates from HumRRO might engage LEARN representatives more and could increase meeting participation.

CHAPTER 3: CAASPP SMARTER BALANCED INTERIM ASSESSMENT ADMINISTRATION STUDY

Caroline Wiley, Wade Buckland, Tom Kiger, and Matthew Swain

Background

The California Assessment of Student Performance and Progress (CAASPP) System of assessments for English language arts/literacy (ELA) and mathematics includes the three components of the Smarter Balanced Assessment System: (a) summative assessments, (b) interim assessments, and (c) the Digital Library of formative assessment processes. Ideally, the three individual components should work together to contribute to the overall purpose of promoting high-quality teaching and learning to prepare students for college and career readiness, as illustrated in Figure 3.1.

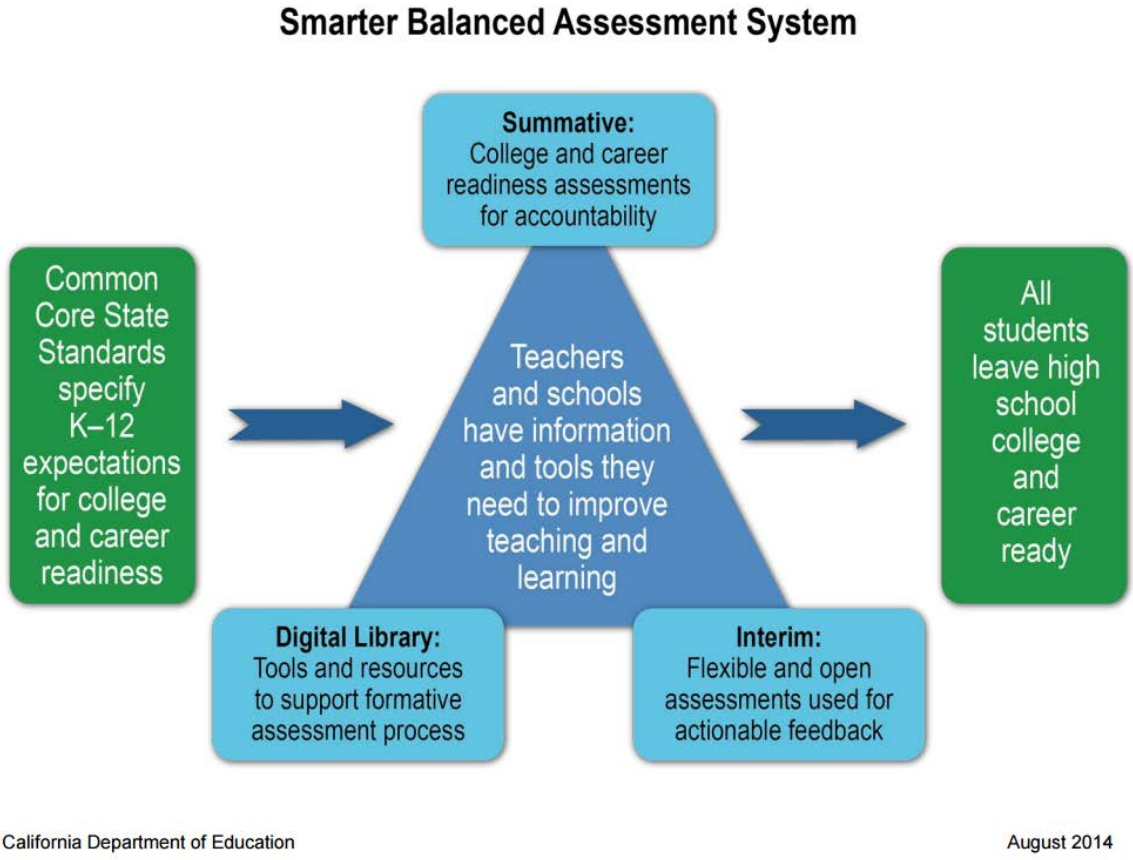


Figure 3.1. Illustration of Smarter Balanced Assessment System.
http://www.cde.ca.gov/ta/tg/sa/documents/sbacgraphic.pdf

Interim assessments, designed to support teaching and learning throughout the school year, play a pivotal role in translating knowledge learned in the classroom to student achievement on the state's end-of-year summative assessments. The interim assessments also provide

information about students' grade level progress towards mastery of the Common Core State Standards (CCSS). The CAASPP System allows local flexibility regarding the Smarter Balanced Interim Assessments, which are optional and were used for the first time in 2015.

To address this knowledge gap and gain insight into this issue, HumRRO's Independent Evaluation Plan called for a CAASPP Smarter Balanced Interim Assessment Administration Study to examine how local educational agencies (LEAs) use these interim assessments and how they administered the interim assessment during the 2015–16 school year.

This chapter presents the design, data sources, and results of the Smarter Balanced Interim Assessment Administration Study. The study relied on two primary sources of data: (a) focus groups and interviews with volunteers from HumRRO's Local Educational Agency Research Network (LEARN[13]), and (b) statewide surveys of LEA CAASPP coordinators, school site CAASPP coordinators, and Smarter Balanced Interim Assessment test administrators. The study's results provide rich and generalizable data on how Smarter Balanced Interim Assessments might be used by LEAs and schools during the 2016–17 school year, including potential successes, challenges, and impacts on instruction.

Smarter Balanced Interim Assessments

A multitude of resources are available to support LEA staff in using the Smarter Balanced Interim Assessments. For example, LEA staff have access to (a) online Webinars, (b) guidance on how to use interim assessments to inform instruction, and (c) technical specifications for accessing the required systems. Because the interim assessments are not state mandated, there is no prescribed guidance for how LEAs should use the interim assessment resources. Rather, LEAs and schools have flexibility in determining the level of support provided for training and administration of the interim assessments. According to the CDE's *Smarter Balanced Interim Assessments* Web site (http://www.cde.ca.gov/ta/tg/sa/sbacinterimassess.asp), the interim assessments are specifically intended to provide:

- Meaningful information for gauging student progress throughout the year toward mastery of the skills measured by the summative assessments; and
- Assessments of the CCSS, which can be used at strategic points during the school year.

There are two types of Smarter Balanced Interim Assessments: Interim Comprehensive Assessments (ICAs) and Interim Assessment Blocks (IABs). Both types are developed for grades three through eight and high school in mathematics and English language arts/literacy (ELA), but can be administered across the full range from kindergarten through twelfth grade. The ICAs are designed similarly to the Smarter Balanced Summative Assessments; that is, they are developed using the same blueprints and item formats, using the same scaling in scoring, and requiring hand scoring. The IABs focus on specific Smarter Balanced targets and are more appropriate for classroom instructional purposes. They too are similar to the summative assessments in item type and format; some IABs require hand scoring.

[13] See chapter 2 for a full description of the LEARN.

Study Design

The purpose of the Smarter Balanced Interim Assessment Administration Study was to explore which interim assessments LEAs decide to use—ICAs, IABs, or both—and their perceptions of those assessments' effectiveness. Because few data were systematically collected about the Smarter Balanced Interim Assessments prior to this study, we planned multiple stages of activities, with each stage informing implementation of the subsequent stage. HumRRO presented the details of the study design at the October 2015 Technical Advisory Group (TAG) meeting. The CDE technical monitor approved the final study design in November 2015, and we commenced work shortly thereafter.

The study began with a focus on knowledge acquisition, involving focus groups and interviews with a small sample of stakeholders from the LEARN, and culminated in statewide data collection using online surveys for the purpose of generalizing results. Study participants were asked to (a) identify problems and strengths with the interim assessments, (b) share their perceptions of how interim assessments may be used to help increase student achievement, and (c) describe the availability of resources needed to administer the interim assessments. To supplement our study data, we obtained statewide Smarter Balanced summative assessment data for the 2016 administration and conducted an analysis to investigate the potential relationship between interim assessments usage and student achievement.

Research Questions

Table 3.1 displays the research questions addressed by this study. For each research question, the table indicates the data sources targeted for our data collection: focus groups and site visit interviews, statewide surveys, or both.

Table 3.1. Research Question Coverage by Data Source

Research Question	Focus Groups and Site Visit Interviews	Statewide Surveys
1. How are decisions made about whether and how interim assessments are used?	✓	✓
2. What are detectable patterns in the types of interim assessments used?	✓	✓
3. To what degree were schools successfully prepared to administer the interim assessments (e.g., training materials were clear, system components worked)?	✓	✓
4. To what degree is the information about test administration procedures, as included in interim assessment resources, followed?		✓
5. To what degree do LEAs perceive the interim assessments impact instructional practice and student achievement?	✓	✓
6. What challenges existed in the 2015-16 school year that could be improved for 2016-17?	✓	✓

Data Sources

To address the designated research questions, we obtained existing CDE data, followed by focus group data and finally statewide survey data. In our original study design, each data source informed the next stage of data collection to examine more comprehensively the processes LEAs used to make decisions about and administer Smarter Balanced Interim Assessments (see Figure 3.2 and Table 3.2). The sample and recruitment of participants for data collection activities is described in a later section.

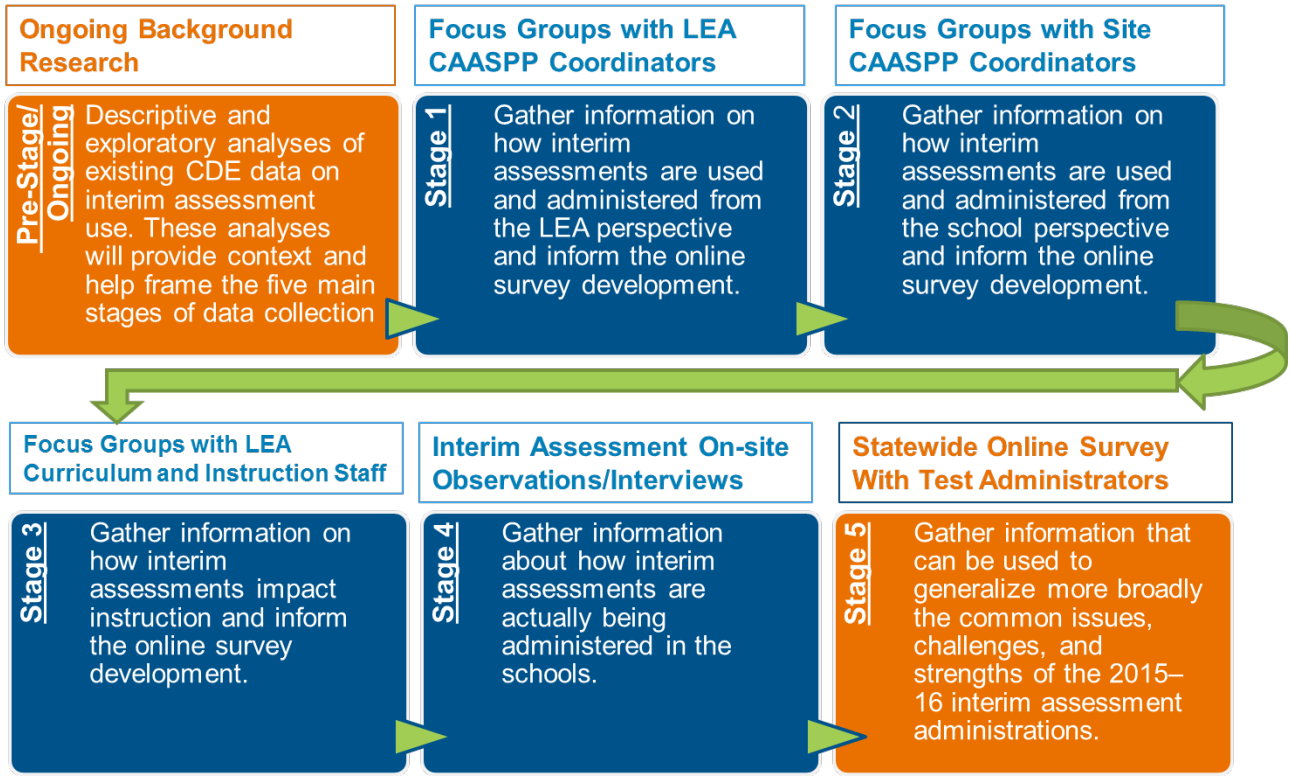


Figure 3.2. Staged design of Interim Assessment Administration Study.

Table 3.2 provides detailed descriptions of information we collected from each data source. Broadly speaking, the data collected during focus groups and site visit interviews were designed to provide additional context surrounding local interim assessment use and were essential to informing development of the statewide survey.

Table 3.2. Interim Assessment Administration Study Data Source Descriptions

Data Source	Description
Existing CDE data on interim assessment use	• Smarter Balanced Interim Assessment usage from the operational contractor. • Descriptive and exploratory analyses of interim assessment usage data by test, grade, LEA, and school during 2015–16 school year. • Analyses provided context and helped frame subsequent data collection efforts.
Focus groups with LEA CAASPP coordinators	• Participants were volunteers from a sample of LEAs from the LEARN. • Gathered information on how interim assessments are used and administered from the LEA perspective. • Each focus group approximately 60 minutes, via teleconference.
Focus groups with site CAASPP coordinators	• Participants were volunteers from schools selected by the LEA CAASPP coordinators. • Gathered information on how interim assessments are used and administered from the school perspective. • Each focus group approximately 60 minutes, via teleconference.
Focus groups with LEA curriculum and instructional staff members	• Participants were volunteers from a sample of LEAs from the LEARN. • Gathered information on how interim assessments impacted instruction. • Each focus group approximately 60 minutes, via teleconference.
School site visit interviews with teachers, site CAASPP coordinators, and test administrators[1]	• Participants were volunteers from one or more schools that HumRRO visited as part of another 2015–16 evaluation activity, the Designated Supports and Accommodations Study. We collaborated with that study. • 10–15 minute in-person interviews that focused specifically on interim assessments.
Statewide online survey[2]	• Participants were respondents to one of three separate surveys targeted to: LEA CAASPP coordinators, school site CAASPP coordinators, and interim assessment test administrators. • Gathered information about common use, issues, challenges, and strengths of the 2015–16 Smarter Balanced Interim Assessment administrations. • Administered in spring 2016 using HumRRO's secure survey platform.

[1] In addition to interviewing test administrators, our original study design included on-site observation of actual Smarter Balanced Interim Assessment administrations. Due to the launch of the study later in the school year and scheduling conflicts with the Smarter Balanced Summative Assessments, HumRRO was not able to observe test administrations as planned.

[2]The original study design included one statewide survey of Smarter Balanced Interim Assessment test administrators. However, based on preliminary focus group findings and the reduced scope of school site visits, HumRRO and CDE determined it was worthwhile to expand the survey to include LEA CAASPP coordinators and school site coordinators.

Study Sample

The study was conducted in two phases. The first phase included activities related to the focus groups and interviews; the second phase included activities related to the statewide survey. Here, we provide a general overview of the complete study sample; additional details about participants in each phase are presented in subsequent sections.

As described in chapter 2, we obtained commitments from 24 LEAs to join the LEARN, which was designed to function as a collaborative network on various aspects of the evaluation studies including participation in data collection activities. During the inaugural LEARN meeting in December 2015, we introduced the Interim Assessment Administration Study and began the recruitment effort. To maximize LEA and school participation in focus groups, we used the following multistep, structured recruitment process:

1. HumRRO provided LEARN representatives with a two-page flier that described the Smarter Balanced Interim Assessment Administration Study, which they could share with their local agencies to encourage study participation (see Appendix C1).
2. HumRRO worked with the LEARN to identify LEAs to participate in focus groups. For focus groups with LEA curriculum and instructional staff, we targeted LEAs that acknowledged the interim assessments had *some* impact on instruction in ways that would provide insight on the utility of the assessments.
3. LEAs could commit to participate in all types of focus groups or any combination. We conducted the first focus groups with LEA coordinators to obtain a broad perspective on local usage. The LEA coordinators then provided recommendations for the types of staff who should participate in other focus groups.

The outcome of the recruitment was that focus groups were conducted with a small sample of LEAs from the LEARN and a small sample of schools from within those LEAs. To the extent possible, the sample included LEAs and schools that administered the interim assessment in various ways (e.g., ICAs, IABs, hand scoring component) and reported having varying levels of technical resources (e.g., sufficient computer access).

Although the study was originally intended to focus on how the Smarter Balanced Interim Assessments impacted classroom instruction and decision making about students, it became evident that few, if any, LEAs in the LEARN actually used the assessments in these ways; many LEARN members did not use the Smarter Balanced Interim Assessments at all.[14] Participation in focus groups required some level of interim assessment usage, while participation in the statewide surveys included a section for respondents who did not administer any interim assessments.

For the survey phase of the study, sampling for the three statewide surveys (LEA CAASPP Coordinator [LEA], school site CAASPP coordinator [SITE], and Test Administrator [TA]) began in April 2016. We included all LEA CAASPP coordinators in the sample, but we selected samples of (a) school site coordinators and (b) TOMS-credentialed interim assessment test

[14] The Smarter Balanced Interim Assessment usage file HumRRO received in December 2015, indicated that some LEAs in the state (outside of the LEARN) were using the Interim Assessments quite heavily; however, many LEARN LEA CAASPP coordinators indicated that interim assessments were not being widely administered in their LEAs.

administrators[15] who administered Smarter Balanced interim tests in the state of California during the 2015–16 school year.[16]

Focus Groups and Interviews

The following sections describe the procedures and findings for the telephonic focus groups and in-person interviews at school sites.

Sample

Table 3.3 summarizes the sample sizes for the focus groups and site visit interviews. Overall, 10 LEAs participated in these activities. The telephonic focus groups consisted of 18 staff members from 8 LEAs. It is important to note that the focus group and interview sample was intended to be small; our purpose was to use these findings to help inform the subsequent statewide survey development rather than to generalize across the state of California.

Table 3.3. Smarter Balanced Interim Assessment Study LEA Focus Group and Interview Samples

ID	LEA Grade-Level	Focus Groups			Site Visit Interviews
		LEA CAASPP Coordinators	Site CAASPP Coordinators	LEA Curriculum & Instructional Staff	Teachers, Site CAASPP Coordinators
1	K-12	Yes	Yes	-	-
2	7-12	Yes	Yes	Yes	Yes
3	PreK-12	Yes	Yes	Yes	Yes
4	K-12	-	Yes	Yes	-
5	6-12 (Charter)	Yes	-	-	
6	K-12	Yes	-	-	
7	PreK-12 (County Office of Education)	Yes	-	-	-
8	PreK-12	Yes	-	-	-
9	7-12	-	-	-	Yes
10	K-8	-	-	-	Yes
	Total N (Participants)	7	6	5	8 schools in 4 LEAs
	Target N (Participants)	8.	12.	8	6-8 schools in 2 LEAs
	Total Focus Groups or Interviews	3 FGs/ 2 Interviews	2 FGs/ 1 Interview	2 FGs	16 Interviews

[15] A TOMS-credentialed interim assessment test administrator is an educator who is registered in TOMS and has a ROLE that is permitted to administer the Smarter Balanced Interim Assessments.
[16] We randomly selected school site coordinators and every 15th credentialed interim assessment test administrator listed on the ETS database of logins for educators allowed to administer the interim assessments.

Protocol Development and Staff Training

The goals of the focus groups were to gather perspectives from local staff on how they use and administer the Smarter Balanced Interim Assessments, and to apply what we learned to the development of items for the statewide survey. We developed five semi-structured protocols that addressed topics aligned to the research questions listed in Table 3.1. Focus group questions and probes were informed by our background research, including knowledge gleaned from our participation in CAASPP institutes and clinics, conversations with LEARN member representatives, review of the online Smarter Balanced Interim Assessment and CAASPP resources, and conversations with CDE staff. Additionally, we conducted background research on prior interim assessment usage by each LEA and school represented in our focus groups, to provide contextual information during data collection.

The questions in each focus group protocol were tailored to the participant type. Specifically, the topics we considered most salient to the LEA CAASPP coordinators and site CAASPP coordinators included (a) the use of the interim assessments and (b) preparation and training. The LEA curriculum and instruction staff focus groups' topics included (a) the use of the interim assessments, (b) impact on curriculum and instruction, and (c) challenges. The site visit interviews, which involved site CAASPP coordinators, test administrators, and teachers, included questions on (a) use of the interim assessments, (b) impact on curriculum and instruction, and (c) challenges. We obtained input from participants about most of the topics. See Appendices C2 through C5 for all focus group and interview protocols.

HumRRO held a session to train internal staff to review the Smarter Balanced and CAASPP Systems and to specifically review the protocols and focus group and interview procedures. The training included relevant background information, an overview of the study design, and a review of the research questions.

Methods

Each focus group was conducted by two HumRRO staff: a facilitator and a note taker who generated detailed summary notes of the focus group. Most interviews were also conducted with two HumRRO staff, although in a few cases the interview facilitator at school sites also served as the note taker. Each detailed summary was reviewed and finalized by staff who participated in the focus group or interview. Emergent themes were identified and coded. Findings in the subsequent section include a discussion of the prominent and salient themes that emerged across focus groups. Although these findings were primarily used to inform development of questions and response options for the statewide surveys, they also provided limited insight into local understanding, use, and perceived value of the Smarter Balanced Interim Assessments.

Focus Group and Interview Findings

The qualitative data collection phase of this study included participants from eight LEAs. Findings from the focus groups and interviews were used to help guide development of the content and phrasing of survey questions and response options. In presenting summaries of our analysis of the qualitative data, our aim is not to imply generalizability of these findings across the state but rather to illustrate the range of viewpoints about the Smarter Balanced Interim Assessments expressed during in-depth discussions with LEA CAASPP coordinators, site coordinators, and teachers. Given our purpose, we did not poll participants to determine how many agreed with specific comments. The summaries also make evident the variation among participants' familiarity with and understanding of the different aspects of the interim

assessments–from setting them up to administering and hand scoring them and then interpreting reports. Some concerns that arose in focus groups and interviews may be the result of LEA-specific differences in readiness for implementing the interim assessment component of the CAASPP System, differences that are also known to exist among LEAs across the state.

Table 3.4 provides a summary of common themes expressed by study participants during focus groups and site visit interviews. The themes are organized across four main topic areas: (a) use of assessments, (b) preparation and training, (c) challenges, and (d) overall. Ultimately, most focus group participants and interviewees did not question the quality of the interim assessment content, but rather raised concern with the usefulness of the interim assessment results. Participants also indicated they understood the interim assessments to be a work in progress and planned to continue using them given continued improvement. A brief discussion of the common themes for each topic follows the table. Within each topic discussion, we provide examples of common as well as insightful but uncommon responses by participant type. It is important to note that our data come from self-reported perceptions of members of the three stakeholder groups and therefore reflect stakeholders' differences in interim assessment usage and experiences.

Table 3.4. Summary of Common Themes in Interim Assessment Study

Topics	Common Themes
Use of Assessments	• IABs were more commonly used than ICAs. • Use of the assessments was typically voluntary (classroom or school decision) but strongly encouraged. • The shorter testing time, absence of hand scoring, and flexibility of IABs was preferred over ICAs. • The interim assessments helped both teachers and students gain familiarity with the Smarter Balanced testing platform. • Many LEAs used third-party assessment platforms in addition to the Smarter Balanced Interim Assessments (e.g., Illuminate Data & Assessment™)
Preparation and Training	• Most reported they received the training they needed. • Several noted they had little training on some topics. • District resources were sometimes seen as more useful than those provided by CDE or Smarter Balanced.
Impact of Assessments	• Feedback in score reports was too broad and had little direct impact on instruction. • Incorporating student interim assessment data into professional development and training was a goal but it was seen as a challenge.
Challenges	• Lack of integration with third-party interim assessment systems. • Scheduling and lack of access to computers, particularly once the summative assessment window opened. • Lack of funding, time constraints, and lack of training for hand scoring. • Minor technological issues were common but were resolved with minor delays. • Inability for LEA coordinators to aggregate data.
Overall	• Most indicated they would use the interim assessments next year. • Most indicated the administration went more smoothly this year (2015–16) than last year (2014–15).

Use of Interim Assessments

Participants reported mixed local policies on mandatory versus voluntary administration. In several cases, use of the interim assessments was voluntary but often strongly encouraged. In a few cases, use of the interim assessments was mandatory, particularly for ICAs. Overall, IABs were said to be more commonly used than ICAs due to the shorter testing time and lack of the hand scoring requirement in most blocks.

Some LEA CAASPP coordinators noted the interim assessments serve to familiarize students and teachers with the format of the summative assessments, essentially acting as a "practice test." CAASPP site coordinators mentioned gaining familiarity with assessments was a key goal of using interim assessments. Specifically, they noted the use of interim assessments allowed students to gain necessary computer skills and increase familiarity with test content and the testing environment. Several site visit interviewees indicated using the interim assessments was an LEA requirement in their district. They indicated a key impact of the interim assessments was for students and teachers to gain familiarity with the assessments, including how to navigate and log into the testing systems. In particular, one interviewee noted use of the interim assessments helped identify which teachers needed more training before administration of the summative assessments. Additionally, interviewees mentioned the interim assessments helped relieve student anxiety on the summative assessments.

Preparation and Training

Views on training and the extent to which training was provided were mixed. Most participants indicated they received the training and resources they needed; however, some contradictions and gaps in training were mentioned. Many reported finding hand scoring challenging and the Smarter Balanced Digital Library confusing and difficult to use. Some confusion about the purpose of the Digital Library and its role in the broader interim assessment system also were reported.

LEA CAASPP coordinators reported the CDE training resources were the most effective training tool. Some cited the online training videos as helpful. Other training sessions that coordinators attended (e.g., LEA training, in-person clinics) were considered by some to be not as helpful. Across CAASPP site coordinators, the amount and types of training varied greatly, with some using very few CAASPP resources. Overall, CAASPP site coordinators reported they felt sufficiently prepared to administer the interim assessments, though several noted their district provided more resources and training than those available from CAASPP and others mentioned they received little training on how to use the interim assessments. LEA Curriculum & Instruction staff thought the training they received was helpful, specifically noting that the training for the Digital Library and field training they received was especially useful to provide teachers examples of expectations for writing items and tasks.

Impacts on Curriculum and Instruction

Comments about the limited helpfulness of interim assessments to inform instruction typically related to the content of score reports. Major concerns were raised regarding the lack of detailed feedback provided in the reporting.[17] Due to the broad nature of the feedback (e.g., score level, reporting categories), most participants believed the interim assessment had very

[17] Due to differences in test blueprints and psychometric limitations, different levels of detail are included in score reports for ICAs and IABs.

little impact on instruction. However, participants noted the interim assessments served a useful role in preparing both students and teachers for the summative assessments. Some Curriculum & Instruction staff indicated they try to incorporate interim assessment data into professional development activities but logistical concerns make this challenging. In addition, several teachers reported using the interim assessment as an instructional tool by working on answers with a whole class.

LEA CAASPP coordinators noted the interim assessments should provide targeted feedback, but teachers did not find the "Below, "At," or "Exceeds the Standard" information sufficiently specific to guide instructional interventions[18]. Similar to the LEA coordinators, CAASPP site coordinators noted the interim assessment did not provide detailed results to inform instruction, and the lack of actionable data in the reporting was disappointing to them. One site coordinator noted a preference to use the Smarter Balanced Interim Assessments (over their currently used third-party assessment system) because they align with the standards, but cited the lack of actionable data as a deterrent.

LEA Curriculum & Instruction staff also indicated the interim assessments had minimal impact on curriculum and instruction and noted that teachers often struggle to understand the score reporting. In particular, educators noted the data from interim assessments are not sufficiently granular, leaving them to struggle with how to interpret the data and respond appropriately. Specifically, they noted target-level data would be useful and that training on how to use these data would be beneficial. One participant indicated there were significant delays with the scoring process in their LEA due to local confusion about that process. Further, some schools did not complete the interim assessments until February, making it difficult for them to have an impact on instruction. Another related issue is that LEA staff does not have access to data or reports at the central office.[19]

Site visit interviewees reported interim assessments impacted instruction by helping them recognize which areas needed further review. At one school, educational apps and programs are being incorporated into the curriculum as a way for students to practice the keyboarding and other computer skills required for the summative assessments. Interviewees at two visited schools mentioned the goal of identifying gaps in instruction and indicated they used backwards mapping of the interim assessments to achieve this goal. In contrast, another school mentioned the lack of detailed score information prevented them from knowing which problems students missed so they could better identify students' knowledge gaps. Similarly, another interviewee noted they previously assessed standards in isolation so it was easy to see if the standard was met or not. With the interim assessments, mathematics standards are connected and build upon each other, so they cannot be assessed independently. This was seen as a beneficial change but one that makes reporting and understanding scores more difficult.

Challenges

In addition to the reporting concerns described relative to impacts on instruction, participants identified logistical challenges in terms of scheduling and allocating time and equipment resources, as well as challenges acquiring resources for hand scoring. Software-related issues

[18] Note that the level of reporting is determined based on the validity of the data that can be reported. That is, granular scores based on few test items are deliberately omitted because they may promote unsupportable teacher actions.

[19] It should be noted that this limitation is an intentional built-in feature of the CAASPP System; LEA CAASPP Coordinators determine who has access to the data.

were also noted as a concern; however, most issues described were relatively minor and were resolved after some delays to testing. Most felt the testing process went more smoothly in 2015–16 than in 2014–15 and that overall, the interim assessments functioned well.

In LEAs that conducted hand scoring, some LEA CAASPP coordinators reported that some scorers found hand scoring to be beneficial in familiarizing themselves with the assessment and with scoring student work and enjoyed the collaboration of the calibration exercises. However, in three of the focus groups, LEA CAASPP Coordinators noted some concerns with hand scoring, including high costs to bring in outside scorers,[20] a cumbersome scoring system interface, and the perception that the task was generally overwhelming (particularly in the beginning). One coordinator noted that the LEA discouraged interim assessments that required hand scoring because of concerns with over-testing.

Several individuals also criticized the assessment data management systems. Specifically, they said the Online Reporting System (ORS) reports had limited capabilities relative to their currently used systems and their reporting needs. An individual from a large LEA stated ORS was used to download data files for subsequent use in the LEA's own student data management system. The Test Operations Management Systems (TOMS) posed some difficulty with assigning roles to staff. The CAASPP test administration system went down occasionally during testing and required inconvenient maintenance schedules. A last area of challenge noted was usability issues with the Digital Library, which some participants described variously as having a "clunky" interface, antiquated navigation, and a lack of resources that connect the assessments to the classroom.[21] The need for separate logins for many of these systems was challenging as several teachers forgot their logins; however, this problem was said to be reduced from last year.[22]

HumRRO is aware that many online resources such as videos, rubrics, and exemplars are available to support hand scoring, including resources developed by the Smarter Balanced Consortium, the CDE, and ETS. Our focus groups were an opportunity to learn how acquainted with these resources some members in the field are, and what ranges of use of the resources they would describe. CAASPP site coordinators' views on hand scoring were mixed, with several site coordinators noting that hand scoring was cumbersome and costly. In particular, one site coordinator reported teachers did not feel as if they adequately understood what a good response looked like; however, the Hand Scoring Training does include model responses for teachers to reference. Another site coordinator noted hand scoring was a bit confusing at first but fairly straightforward after training. Some contradictions about self-reported perceptions and factual knowledge emerged. For example, one site coordinator stated having sample responses for hand scoring would have been helpful; however, exemplars were available in the Interim Assessment Hand Scoring System. This illustrates the difficulty some educators have finding or learning where to find specific resources. Many site coordinators experienced some technical glitches with the delivery system, causing delays in starting the exams and resulting in students

[20]The hand scoring requirement for ICAs and some IABs was designed to encourage professional learning for educators in the LEAs, because the interim assessment items that require hand scoring are similar to the hand-scored items in the summative assessments.
[21] It is unknown whether participants' experiences were based on the original or modified Digital Library landing page; however, the data collection occurred during 2016, after improvements had been made.
[22] The burden of two separate logins was somewhat reduced after an individual's username and password were allowed to be the same for TOMS and for the Smarter Balanced Digital Library.

being logged out of exams.[23] One site coordinator noted it was problematic that both the interim and summative assessments were available from the same location as one teacher mistakenly administered a summative assessment. It should be noted that although this was addressed by the operational contractor for the 2015–16 school year to help reduce test selection errors, the coordinator still noted it as an issue. One teacher noted concern with the designated supports in the delivery platform, specifically the choppy and delayed American Sign Language (ASL) interpretation video.

While empirical data regarding the volume and nature of technical support for interim assessments sought by LEA CAASPP Coordinators could be obtained by the CDE from ETS using records from the California Technical Assistance Center (CalTAC), our focus groups were an opportunity to learn about the types of problems that could be important to capture in our survey response options. We captured what our interviewees said, but because we have no way of distinguishing at which point in time they reference (e.g., before an upgrade eliminated their concern), we did not examine whether their attributions about the causes of their concerns were correct. LEA curriculum and instruction staff noted several challenges with the use of interim assessments related to technology and logistics. For hand scoring in particular, they had trouble finding scorers, securing funding for scorers, and completing the large volume of work in a timely manner. Several noted the CAASPP Web sites were not always accessible due to Web site closures or maintenance, and there were difficulties with passwords. Further, one LEA staff member stated the Web site went down during an assessment. LEA staff noted difficulties arranging student access to devices and logging onto the assessment platform.

Site visit interviewees at several schools reported experiencing difficulty with hand scoring. One school indicated ELA was especially challenging because teachers were used to grading multiple categories (including spelling) rather than only one skill.[24] Another school noted the inability to sort student responses in the Interim Assessment Hand Scoring System as an obstacle that limited how responses could be assigned to scorers, which slowed the pace of scoring. In contrast, one school noted the hand scoring workshop was considered helpful from a professional development perspective and teachers were pleased with its value. Many schools indicated there were some minor technical difficulties such as issues with student login and system crashes, but these were resolved with only short delays to testing. Several schools noted that using interim assessments was beneficial in determining what technical difficulties to address and which teachers needed more training and support prior to the summative assessment.

Statewide Survey Methods

The statewide survey was originally designed as a broad examination of test administrator experiences related to administering the interim assessments. However, based on preliminary focus group findings and the scheduling conflicts that prevented us from conducting the interim test administration observations, both HumRRO and CDE determined it was worthwhile to expand the survey to also gather input from all LEA CAASPP coordinators and a systematic random sample of school site coordinators. The sections below describe the sampling, survey development, and analytical procedures and findings.

[23] The focus group discussion did not deeply investigate the root causes of the noted technical issues. Therefore, it is uncertain whether the issues were with the delivery system itself, with the local internet capacity, or some other related cause.

[24] Many Interim Assessments also include multiple rubrics and some even contain conventions. The local hand-scoring requirement of some IAs was designed from the outset to promote professional development for educators in LEAs.

Sampling Procedures

To draw the final samples for each survey (LEA, SITE, and TA), we used the CDE District Coordinator contact list (for LEA) and TOMS user role data exports (for SITE and TA). We created sampling frames for each survey using standard data cleaning procedures (e.g., removing duplicate coordinators and administrators, removing users with no associated County-District-School [CDS] code). Table 3.5 describes the general sampling process and final sample counts for each survey.

As a general note, when referring to the three survey types, we use *LEA, SITE,* and *TA.* When referring to specific participants, we will use *LEA CAASPP coordinator, site CAASPP coordinator,* and *test administrator.*

Table 3.5. Sampling Frame Descriptions for Each Survey

Survey	Sampling Frame Description	Final Count
LEA	• Data source was CDE District Coordinator Contact List • Selected all LEA CAASPP Coordinators from active LEAs (districts and charters)	N=1,599
SITE	• Data source was ETS TOMS User Role data extract (User Role = Test Site Coordinator) • Randomly selected a single school if multiple schools per coordinator e-mail • Randomly selected a single user if multiple coordinator e-mails per school • Randomly selected 1,006 records.	N=1,006
TA	• Data source was ETS TOMS User Role data extract (User Role = Test Administrator, Test Examiner, Test Administrator & Test Examiner, IA Administrator Only) • For the IA Administrator Only role, excluded users associated with more than one LEA and more than one school • For the Test Administrator, Test Examiner, and Test Administrator & Test Examiner roles combined, excluded users associated with more than one LEA and more than one school • If duplicate user e-mails existed, selected the IA Administrator Only role record • Sorted by school and role and selected every 15th record	N=12,751

Survey Development

HumRRO used information from the focus groups as well as usage data on the Smarter Balanced Interim Assessments to develop three surveys that were administered using HumRRO's Survey Platform. The surveys included several sections that focused on (a) usage, (b) assessment administration, (c) accommodations, (d) impacts on instruction, (e) preparation and training, (f) challenges and improvements, (g) next year's plans, and (h) background characteristics.

For each survey, the first question established respondent eligibility (e.g., "Did you administer a Smarter Balanced Interim Assessment this school year?"). Respondents who administered the Smarter Balanced Interim Assessments were routed to a series of questions about their experience; this survey took approximately 15–20 minutes to complete. Respondents who did

not administer any Smarter Balanced Interim Assessments were routed to a shorter section on non-usage; this survey took approximately 5–10 minutes to complete.

Each survey included questions addressing topics related to the Smarter Balanced Interim Assessment and was customized to the point of view and role of the nominee type. Figures 3.3 through 3.5 indicate the topics included in each survey; the LEA survey did not include questions on Assessment Administration or Accommodations as those were more appropriate for the school-level educators. Many questions were implicitly the same across each of the surveys; however, language was appropriately changed to address each respective sample (e.g., Did your LEA…? vs. Did your school…?).

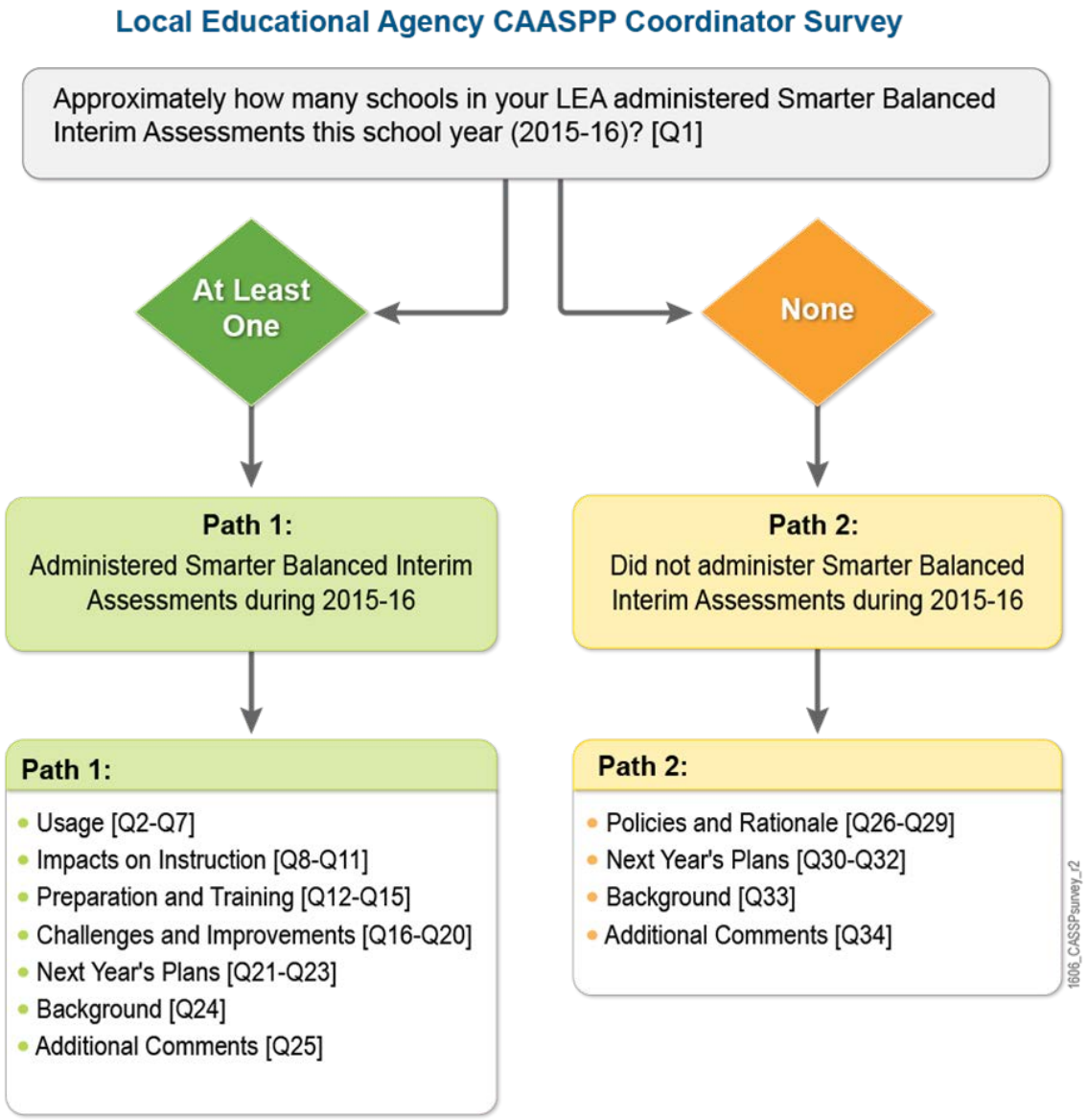


Figure 3.3. LEA CAASPP coordinator survey blueprint.

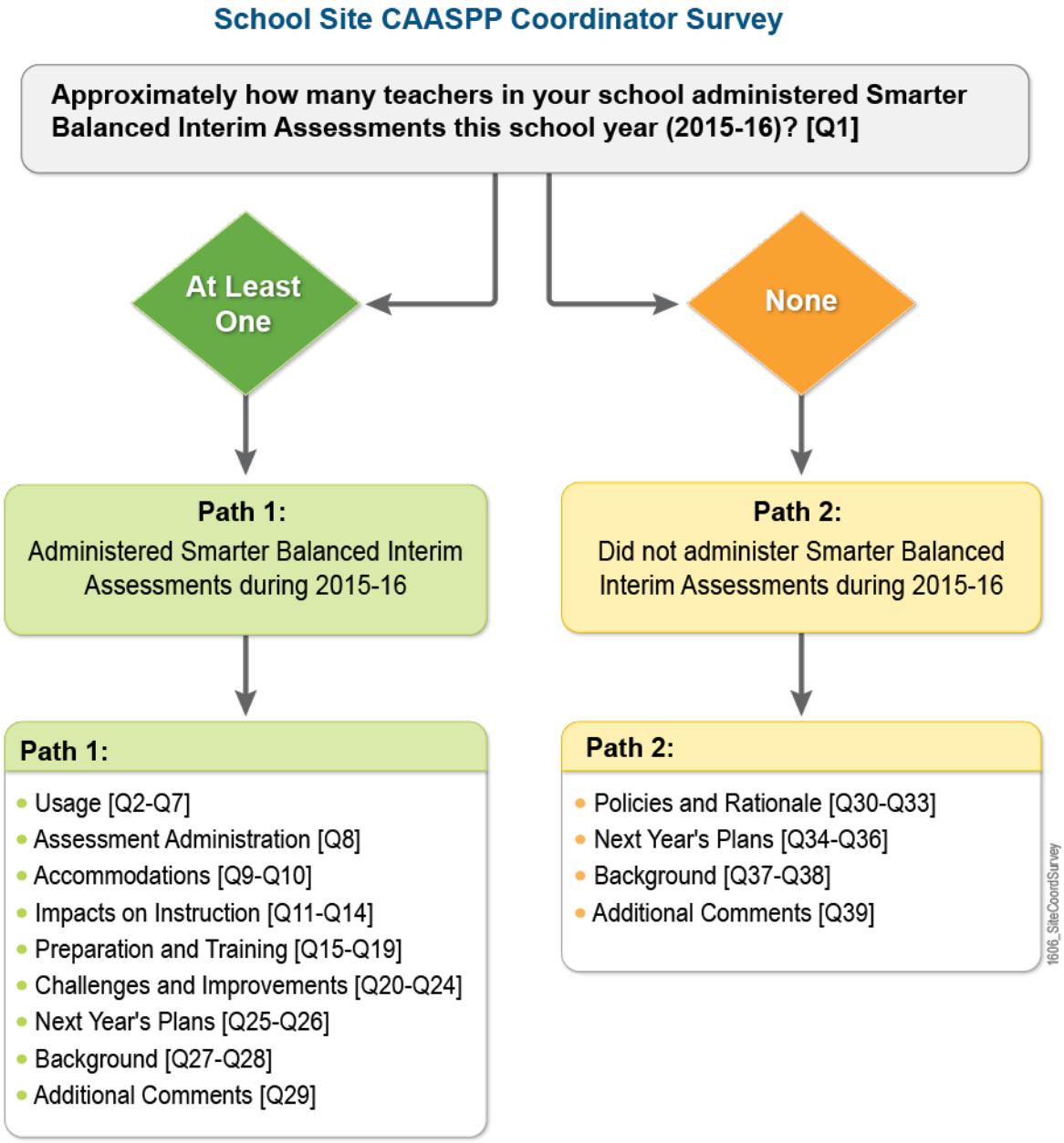


Figure 3.4. School site CAASPP coordinator survey blueprint.

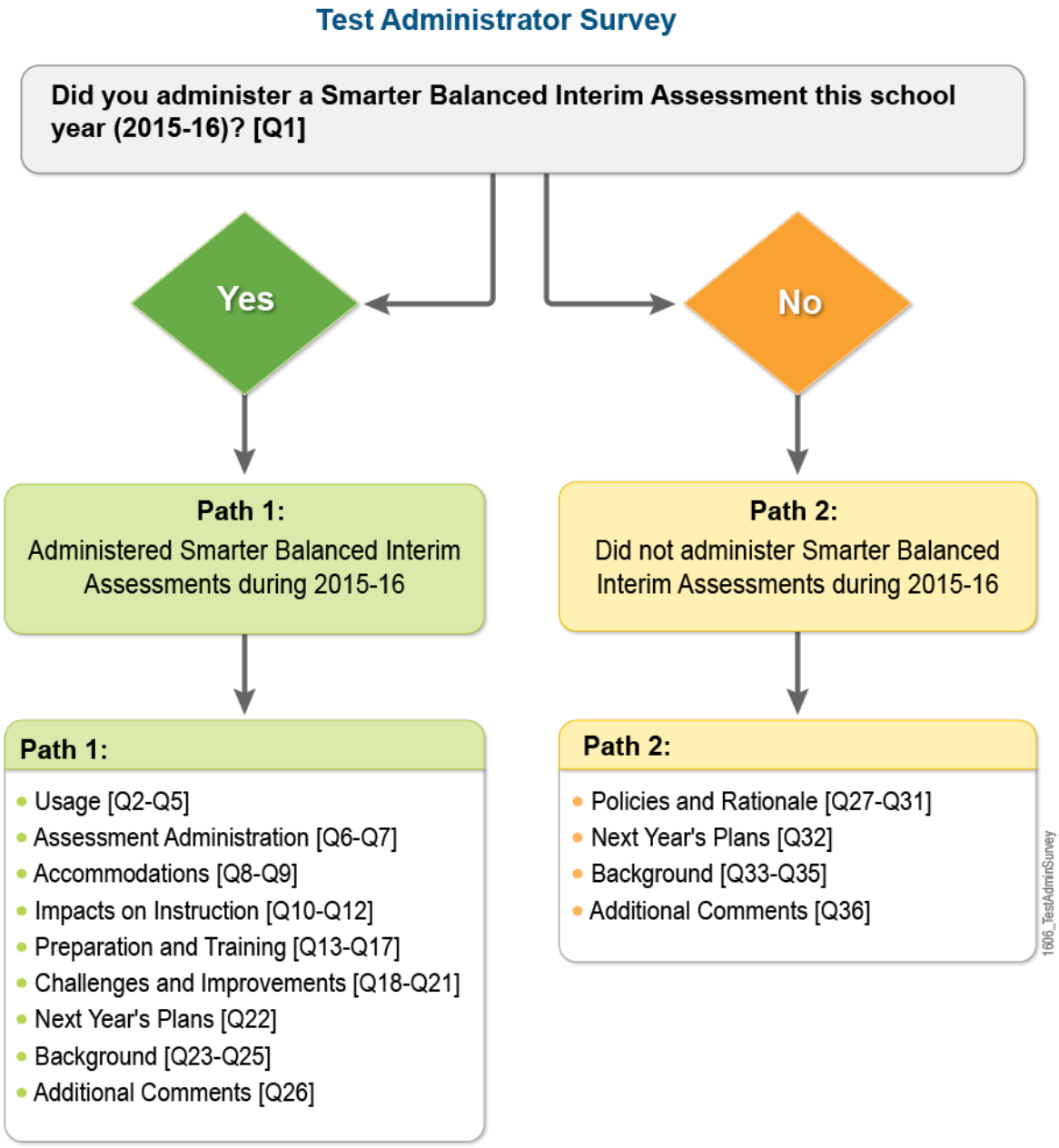


Figure 3.5. Test administrator survey blueprint.

The total number of questions on a given survey ranged from 34 questions for the LEA CAASPP coordinators to 39 questions for the school site CAASPP coordinators. Although most question formats were either *selected response* or *select all that apply,* each survey included a small number of Likert-scale rating items (presented as tables) and short open-ended questions. The three survey blueprints are presented in Figures 3.3–3.5, including the branching logic. Following initial development, HumRRO submitted draft versions of the instruments as Microsoft Word files to CDE for review and approval with particular attention to:

- verifying the survey language aligned with California's intent and the language of the Smarter Balanced Interim Assessments;

- suggesting additional questions to ask respondents; and
- identifying questions considered not appropriate or of questionable value to our research goals.

CDE provided minimal revisions to clarify information provided in the survey. After the survey content was finalized, HumRRO provided access to the Web versions of the instruments for the CDE technical monitor's review and approval. Appendices C6 through C8 present the three final surveys, including all items and response options.

Survey Administration

After survey programming was complete, we deployed a series of activities to ensure a successful launch that would achieve maximal response rates. A number of prelaunch activities are described below. The survey window remained open for approximately two weeks from May 11, 2016 until May 27, 2016. During this time, HumRRO staff responded to telephone and e-mail inquiries about the study, including requests for the survey links to be resent, updates to recipient e-mail addresses, and technical assistance with the survey instrument itself. The following activities were completed to administer the surveys.

1. The CDE e-mailed LEA CAASPP coordinators about one week before the survey launched to endorse it, provide general information about the study, and introduced HumRRO as the independent evaluator (see Appendix C9). Additionally, the message asked LEA CAASPP coordinators to forward the Uniform Resource Locator (URL) and e-mail specifications so LEA information technology staff members could ensure the e-mails were actually delivered and the survey Web page could be opened. CDE also announced the survey in two of the weekly CAASPP update e-mail newsletters, encouraging educators to respond.
2. HumRRO used a delivery program to send e-mails and track respondent behavior (e.g., how many opened the link, how many e-mails bounced). HumRRO sent initial survey invitations, including information about confidentiality of the responses and each user's unique survey link (see Appendix C10).
3. HumRRO resolved, where possible, e-mail addresses that were invalid due to inaccuracies in the address itself (e.g., spelling error of recipient name). We worked with the IT coordinator at several LEAs to resolve blocked delivery due to filtering.
4. Two reminder e-mails were sent to those who had not yet responded.
5. During the last days of the survey window, HumRRO identified LEA CAASPP coordinators who (a) had not yet completed the survey, (b) had educators who were sent invitations for the site coordinator and TA surveys, and (c) were from LEAs with relatively high interim assessment usage. HumRRO called 144 LEA CAASPP coordinators to remind them to complete the survey and to encourage them to endorse the survey to others within their LEA.

Security of Collected Data

We provided user authentication using a unique link for each respondent. Each survey was hosted on HumRRO's secure server. No personally identifiable information (PII) was collected; individual school names and school personnel were not reported. We aggregated results to avoid identification of a particular school by examination of the data.

Response Rates and Respondent Characteristics

Response Rates

In evaluating the response rate of each survey, we analyzed the number of respondents as well as the type of respondent behavior (e.g., whether respondent opened the survey link at all; respondent opened the link, but didn't answer any items). Table 3.6 shows the response rates and behaviors for each survey by (a) sample, (b) response rates, and (c) completion rates.

The Total Delivered rates are the number of participants (out of the number of participants who were sent the survey invitation) who actually received the initial e-mail survey invitation. Reasons for unsuccessful delivery included invalid e-mail addresses and bounced e-mails due to recipient firewall settings. The large majority of undelivered surveys was due to bounced addresses versus invalid e-mail addresses. If an e-mail from a LEARN LEA or a high-volume interim assessment LEA bounced addresses, we resent the invitations manually from the project e-mail.[25] The resulting number of successfully delivered survey invitations was 14,911 (LEA = 1,578; SITE = 986; TA = 12,347).

Table 3.6. Survey Response Rates and Respondent Behaviors

Behavior	LEA n (%)	SITE n (%)	TA n (%)
A. Sample			
Number Sampled	1,599	1,006	12,751
Undelivered	23 (1.4%)	22 (2.2%)	517 (4.1%)
Total Delivered	1,578 (98.7%)	986 (98.0%)	12,347 (96.8%)
B. Response Rates of Total Delivered			
Opened survey	454 (28.8%)	165 (16.7%)	1,683 (13.6%)
Opened survey, answered at least one item	385 (24.4%)	137 (13.9%)	1,303 (10.6%)
Opened survey, answered at least 90% of items	336 (21.2%)	82 (8.3%)	738 (6.0%)
Opened survey, didn't answer any items	69 (4.4%)	28 (2.8%)	380 (3.1%)
C. Survey Completion Rates			
Mean % items answered (of those who answered at least one item)	91.1%	68.2%	65.0%
Respondents deleted due to low completion rate[a]	18	30	259
Total analytic sample (of the Total Delivered)	367 (23.2%)	107 (10.1%)	1,044 (8.5%)

[a] Thresholds for removing low completers are in Table 3.7.

[25] HumRRO used MailChimp (a Web-based e-mail marketing service) to send out initial e-mail invitations and reminders. A project e-mail address was used for survey correspondence with participants.

Table 3.6 details the percentages of respondents who at least opened the survey (clicked on the survey link and saw the landing page), the percentage who continued on to answer at least one item, and the percentage who persisted to answer at least 90 percent of the items. Small percentages of participants opened the survey, but did not answer any of the items (LEA = 4.4%; SITE = 2.8%; TA = 3.1%).

We computed survey completion rates for each survey to determine how many items respondents completed. Because each survey comprised different paths a respondent could take based on their Smarter Balanced Interim Assessment usage, varying numbers of total items were shown on the survey to any given respondent (i.e., some paths included items that other paths did not). Table 3.7 describes the different paths and sub-paths for each survey and the respective cutoffs for exclusion. The final sample used for analysis was 1,518 (LEA = 367; SITE = 107; TA = 1,044). The overall response rate was 14.2 percent, with the rate highest for the LEA coordinators (23.2%), followed by site coordinators (10.1%), and lowest for test administrators (8.5%). Respondents with low completion rates were removed from analyses. The very low response rates, particularly for the test administrators, preclude us from interpreting findings as generalizable statewide; however, the responses from approximately 1,300 educators do suggest a range of experiences with and opinions about the interim assessments.

Table 3.7. Survey Exclusion Criteria by Survey Path

Path	Description	Criteria for Exclusion: Number of Survey Items Completed (% of Total Number of Path Items)
LEA and SITE Surveys		
Path 1A.1	Used IAs, they were required or highly encouraged, and they will NOT permit IAs next year.	LEA: None excluded SITE: No respondents on this path
Path 1A.2	Used IAs, they were required or highly encouraged, and they will permit IAs next year.	LEA: <= 1 item (1.3%) SITE: <= 1 item (1.2%)
Path 1B.1	Used IAs, they were voluntary and not encouraged, and they will NOT permit IAs next year.	No respondents on this path
Path 1B.2	Used IAs, they were voluntary and not encouraged, and they will permit IAs next year.	LEA: <= 4 items (5.2%) SITE: None excluded
Path 2A	Did NOT use IAs, but they will permit them next year.	LEA: <= 2 items (25%) SITE: <= 2 items (19.0%)
Path 2B	Did NOT use IAs, but they will NOT permit them next year.	LEA: None excluded
TA Surveys		
Path 2A	Did NOT use IAs, but they will permit them next year.	<= 2 items (2.2%)
Path 2B	Did NOT use IAs, but they will NOT permit them next year.	<= 2 items (20%)

Respondent Characteristics

We compared statewide LEA demographics with those of the survey sample (Table 3.8). The larger differences between the analytic site and TA samples are primarily explained by drawing a random sample rather than a matched proportional sample and by making recruiting attempts to include LEAs that had used the interim assessments. Note that the LEA and TA samples tended to be from larger districts, with larger percentages of students eligible for free or reduced priced meals (FRPM), and with more ELs. The characteristics of our analytic sample, especially for the site and TA roles, are not broadly representative of the state as a whole and limit our ability to generalize our findings. The survey findings we report should therefore be considered only within the context of our sample.

Table 3.8. Survey LEA Demographics Compared to the State

Demographic	State[a] %	Total Sample[b] %	Analytic Sample[c]		
			LEA %	Site %	TA %
English Learners					
High	52.2	52.2	58.5	84.1	70.1
Low	47.8	47.8	41.5	15.9	29.9
Size					
Large	27.6	27.8	53.4	78.5	79.0
Medium	30.8	31.1	26.3	15.9	17.9
Small	41.6	41.1	20.4	5.61	3.07
Free or Reduced Priced Meal (FRPM) Eligibility					
Low % of FRPM	25.1	25.2	20.7	9.35	17.8
Moderate FRPM	49.9	50.2	51.7	49.5	49.9
High % of FRPM	25.0	24.9	27.6	41.1	32.3
Region[d]					
Central Region	25.6	25.1	20.3	12.2	17.1
Coastal Region	23.5	23.7	21.4	21.5	20.7
Mother Lode Mountain Region	9.5	9.3	10.4	5.6	7.2
Northern Region	17.7	17.8	10.7	0.9	2.9
Southern Region	23.7	23.1	37.3	59.8	51.3

[a] State percentages are based on the CDE Student & School Data Files (http://www.cde.ca.gov/ds/sd/sd/#pub)
[b] Total unique LEAs in the original sample
[c] Includes all records (not just unique LEAs)
[d] Based on 2002 California Department of Social Services Regional Groupings (http://www.dss.cahwnet.gov/research/res/pdf/multireports/RegionsofCalifornia.pdf)

Table 3.9 provides descriptive statistics of the analytic samples for the school site coordinator and test administrator surveys. The 107 school site coordinators represented 76 different LEAs. The 1,044 test administrators represented 376 different LEAs and 981 different schools.

Table 3.9. Descriptive Statistics for School Site Coordinator and Test Administrator Analytic Samples

	Test Administrators		School Site Coordinators	
	N	Average Number of Test Administrators (SD)	N	Average Number of Site Coordinators (SD)
Total Analytic Sample	1,044	-	107	-
LEAs Represented	376	2.8 (5.0)	76	1.4 (1.9)
Schools Represented	981	1.1 (0.3)	-	-

Table 3.10 provides insight into how the LEA CAASPP Coordinator analytic sample relates to the test administrator and school site coordinator samples. Of the 367 LEA coordinators, 174 LEAs had representation in the analytic test administrator survey and 41 had representation in the school site coordinator survey.

Table 3.10. Descriptive Statistics for LEA Coordinator Analytic Sample

LEA Participants	N (%)	Average Number Per LEA (SD)
Total Respondents	367	-
With TA Respondents	174 (47.4%)	3.5 (7.0)
With Site Coordinator Respondents	41 (11.2%)	1.6 (2.5)

Analysis

We conducted descriptive analyses of each survey and thematically coded the open-ended questions. All analyses included only the analytic sample. Appendix C11 provides a map to specific survey question numbers.

Statewide Survey Findings

Findings in this section are organized according to major topics addressed by the focus groups and surveys. Because different sets of questions were included on each type of survey, the tables do not present questions in the same order they were viewed by survey respondents. Although the numbers are occasionally small, we believe these findings provide important illustrations of the LEAs' and schools' use of the Smarter Balanced Interim Assessments. They also highlight the perceptions of current educators as to the assessments' strengths and areas in possible need of improvement.

Because the numbers of respondents per survey type vary, it is critical for the reader to attend to the total counts when interpreting percentage values in the tables. Although there are some common survey items across all three surveys, we present the findings separately for each survey (LEA CAASPP coordinator, school site coordinator, test administrator). Additionally, despite some items having similar stems across each survey, the frame of reference for each group is different (e.g., LEAs were asked to think about their district, teachers were asked to think about their classrooms).

Thus, in some sections below, we discuss findings simultaneously for each survey, but we often discuss each survey separately. The survey findings are organized by topic (e.g., use, administration), and all descriptive statistics for the individual items in each survey are presented in Appendix C12. In the beginning of each section, we provide a reference table that displays the section's survey question numbers and corresponding appendix tables.

Background

Table 3.11 shows the survey question numbers related to respondents' background with their corresponding appendix table numbers for individual item descriptive statistics.[26]

Table 3.11. Survey Background Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q24	5 – 6
SITE	Q27 - Q28	7, 9
TA	Q23 - Q25	8, 11

Table 3.12 shows only the most commonly reported job titles of all respondents for each survey. LEA CAASPP coordinators had the widest variety of job titles, including data analyst, teacher, and superintendent.

Table 3.12. Most Commonly Reported Titles for Each Survey

What best describes your title?			
	Valid N	Frequency	Percent
LEA			
LEA Assessment Coordinator	333	183	55.0
Director of Assessment	333	64	19.2
SITE			
School Testing Coordinator	83	32	38.6
Principal/Assistant Principal	83	30	36.1
TA			
Teacher	731	669	91.5

[26] Appendix C11 is a complete cross-reference table that lists all survey topics and their related question numbers in the three surveys as well as the corresponding appendix table numbers for individual item descriptive statistics.

Interim Assessment Usage

Table 3.13 shows the survey question numbers related to interim assessment usage with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.13. Interim Assessment Usage Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q2 - Q7	17, 20 – 22, 26 – 27, 32
SITE	Q2 - Q7	18, 20, 23 – 24, 28 – 29, 33
TA	Q2 - Q5	19, 20, 25, 30 – 31

Survey Paths

For the LEA and school site CAASPP coordinator surveys, respondents answered questions based on their usage, which resulted in six potentially different paths. Path 1A2 (Smarter Balanced Interim Assessments were required or highly encouraged, and they will permit their usage next school year) was the most frequent response on both surveys (LEA = 64.0% SITE = 81.3%). For the test administrator survey, responses to questions resulted in two different paths (Did not administer and did administer interim assessments). Approximately 80 percent of respondents on each of the three surveys reported administering at least one interim assessment (see Figure 3.6).

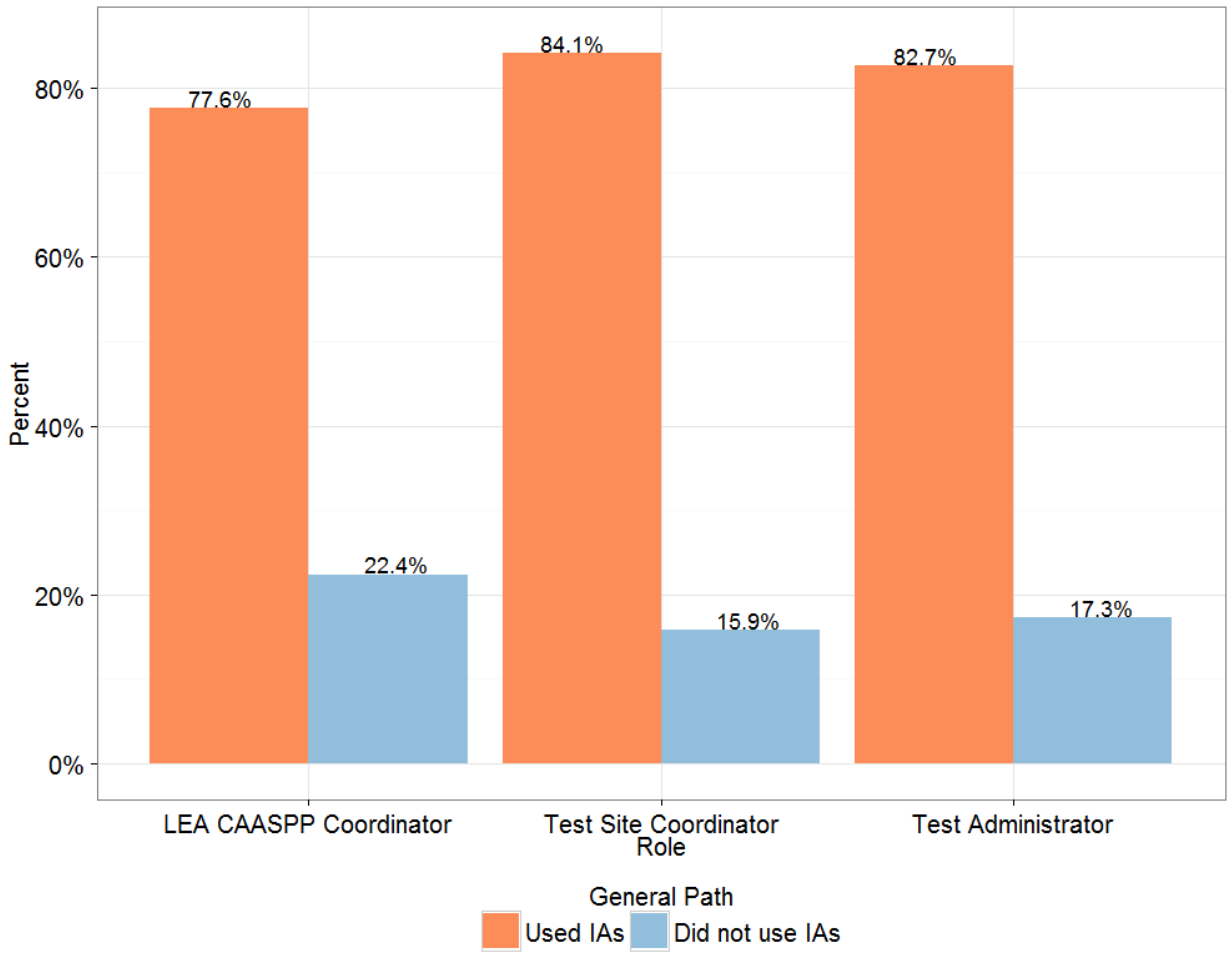


Figure Note: Analytic sample for each survey (Paths 1 and 2): N_{LEA} = 367; N_{SITE} = 107; N_{TA} = 1,044

Figure 3.6. General path for each role (survey).

Figure 3.7 displays Smarter Balanced Interim Assessment usage reported by LEA and school site coordinators. Summing the percentages of coordinators who indicated 'many' and 'most' shows that 50 percent of LEA coordinators stated that at least half of the schools in their LEAs administered the assessments; 46 percent of site coordinators stated that at least half of the teachers in their school administered the assessments. Test administrators were not asked to quantify their usage, but a large majority (83%) of respondents reported administering at least one assessment during the 2015–16 school year.

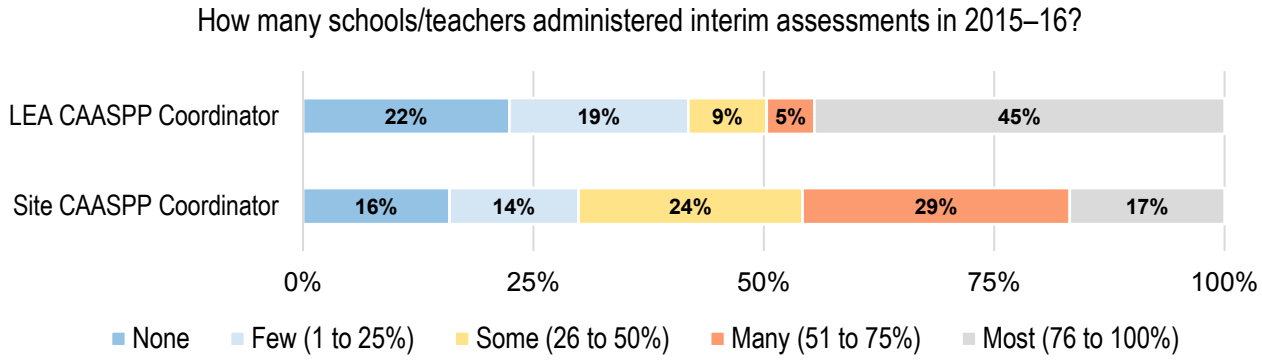


Figure Note: Analytic sample for each survey (Paths 1 and 2): N_{LEA} = 367; N_{SITE} = 107; N_{TA} = 1,044

Figure 3.7. Reported usage of interim assessments (LEA and SITE surveys).

Because we were able to compute the number of interim assessments administered statewide in 2015-16 using data from ETS (see Table 3.14), we did not collect frequencies of interim assessment administrations on our surveys. According to the preliminary usage data from ETS through March 2016, prior to our survey launch, there were just over four million administrations of IABs and ICAs in 2015-16 (combining CAT and PT administrations). As expected, that is an increase from the 2014-15 school year, when interim assessments were only administered starting January 2015. In addition, CDE has made considerable efforts to communicate availability and resources regarding the Smarter Balanced Interim Assessments.

Table 3.14. Count of Interim Assessments Administered During 2014–15 and 2015–16

	ELA		Math		Overall	
Type	2014-15[1]	2015-16	2014-15[1]	2015-16	2014-15[1]	2015-16
IAB	353,613	1,803,281	262,906	1,636,143	616,519	3,439,424
ICA	64,671	457,924	65,421	424,064	130,092	881,988
Total	418,284	2,261,205	328,327	2,060,207	746,611	4,321,412

[1] California first offered the interim assessments to LEAs in late January 2015.
Sources: 2014–15 data are ETS raw Interim Assessment data. 2015–16 data are ETS Interim Assessment aggregate data, through March 2016.

In the subsequent sections, we present results first for respondents who administered at least one interim assessment (Path 1), followed by brief descriptions of those who did not administer any interim assessments.

Policies

Approximately 50 percent of LEA and school site coordinators reported that the use of the Smarter Balanced Interim Assessments was required by their LEA[27] (see Table 3.15). A small percentage of school site coordinators (15%) reported their LEA made usage voluntary while their school required use of the interim assessments.

Table 3.15. Reported Interim Assessment Policies

What best describes your LEA's policy regarding administration of Smarter Balanced Interim Assessments this school year?			
	LEA	SITE	TA
VALID N	279	86	843
The LEA required it for one or more grade/subject combination(s).	134 (48%)	49 (56%)	705 (83%)
Our LEA made it voluntary, but my school required it for one or more grade/subject combination(s).	---	13 (15%)	33 (4%)
It was voluntary, but highly encouraged.	99 (35%)	23 (26%)	89 (10%)
It was voluntary with no specific encouragement.	49 (17%)	3 (3%)	25 (3%)

A much higher number of test administrators reported their LEA required some administration of the interim assessments (83%). Upon further investigation, there is evidence to suggest that test administrators' understanding of required policies does not match that of their LEA coordinators. Table 3.16 indicates that 83.8 percent of test administrators agreed with their LEA coordinators who also reported the interim assessments were a requirement. Conversely, a rather high percentage of test administrators reported the interim assessments were an LEA requirement in LEAs whose coordinators reported they were *not* a requirement (74.6%). These findings suggest a lack of familiarity or misunderstanding between LEA requirements and the classroom-level educators' understanding of those requirements.

[27] School Site Coordinators reported independently of their respective LEAs. That is, this is not an LEA-matched statistic.

Table 3.16. Percentage of Test Administrators Reporting Interim Assessments as a Requirement Matched by LEA

What best describes your LEA's policy regarding administration of Smarter Balanced Interim Assessments this school year? – crossing TA and LEA results			
	LEA Frequency	Avg % of TAs	Avg Number of TAs per LEA
TAs reporting interim assessment is LEA requirement when LEA reported it is a requirement	67	83.8%	2.6
TA reporting interim assessment is LEA requirement when LEA reported it is not a requirement	72	74.6%	3.9

Note: The total number of LEAs reported here is smaller than the total number of LEAs in the LEA Path 1 sample (n=282) because (a) not all LEAs followed path 1, or (b) not all LEAs had TAs included in the analytic sample.

Administered Assessment Types

In LEAs and schools that reported requiring or highly encouraging use of the Smarter Balanced Interim Assessments, the Interim Assessment Blocks (IABs) were the most commonly cited assessment (LEA: Mathematics IAB = 76.0%, ELA IAB = 74.2%; SITE: Mathematics IAB = 79.50%, ELA IAB = 71.10%). The IABs were the most commonly cited assessment type in LEAs that did not provide any specific encouragement (Mathematics IAB = 85.4%, ELA IAB = 81.3%).[28] Test administrators reported administering a mix of assessment types (Mathematics ICA = 59.9%, Mathematics IAB = 44.3%, ELA ICA = 65.1%, ELA IAB = 45.6%)

When examining the different combinations of interim assessment use, similar patterns emerge. Table 3.17 shows that LEA and school site coordinators reported requiring or highly encouraging both the mathematics and ELA IABs, whereas test administrators most frequently reported administering the ICAs and myriad other combinations.

[28] Due to the small number of school site coordinators that reported no specific encouragement of interim assessments, frequencies of this group are presented only in the appendix.

Table 3.17. Most Commonly Reported Combinations of Interim Assessment Type Usage

	Frequency	Percent
LEA		
Mathematics IABs, ELA IABs	101	44.1%
Mathematics ICAs, ELA ICAs, Mathematics IABs, ELA IABs	50	21.8%
Mathematics ICAs, ELA ICAs	33	14.4%
All other combinations	45	19.7%
SITE		
Mathematics IABs, ELA IABs	46	55.4%
Mathematics ICAs, ELA ICAs	14	16.9%
Mathematics ICAs, ELA ICAs, Mathematics IABs, ELA IABs	12	14.5%
All other combinations	11	13.2%
TA		
Mathematics ICAs, ELA ICAs	237	29.2%
Mathematics ICAs, ELA ICAs, Mathematics IABs, ELA IABs	140	17.2%
Mathematics IABs, ELA IABs	124	15.3%
All other combinations	311	38.3%

Primary Uses

Across both LEA and school site coordinators, the most commonly cited uses of the interim assessments that contributed to requiring or highly encouraging their use was to familiarize students and teachers with testing/testing systems (LEA = 93%; SITE = 86%), familiarize students and teachers with item types (LEA = 83%; SITE = 83%), and to provide students practice with universal tools, designated supports, and accommodations (LEA = 78%; 75%). Respondents were asked to select all uses that applied. Very few coordinators reported primary uses to inform subject matter professional development offerings or to determine student placement (see Figure 3.8).

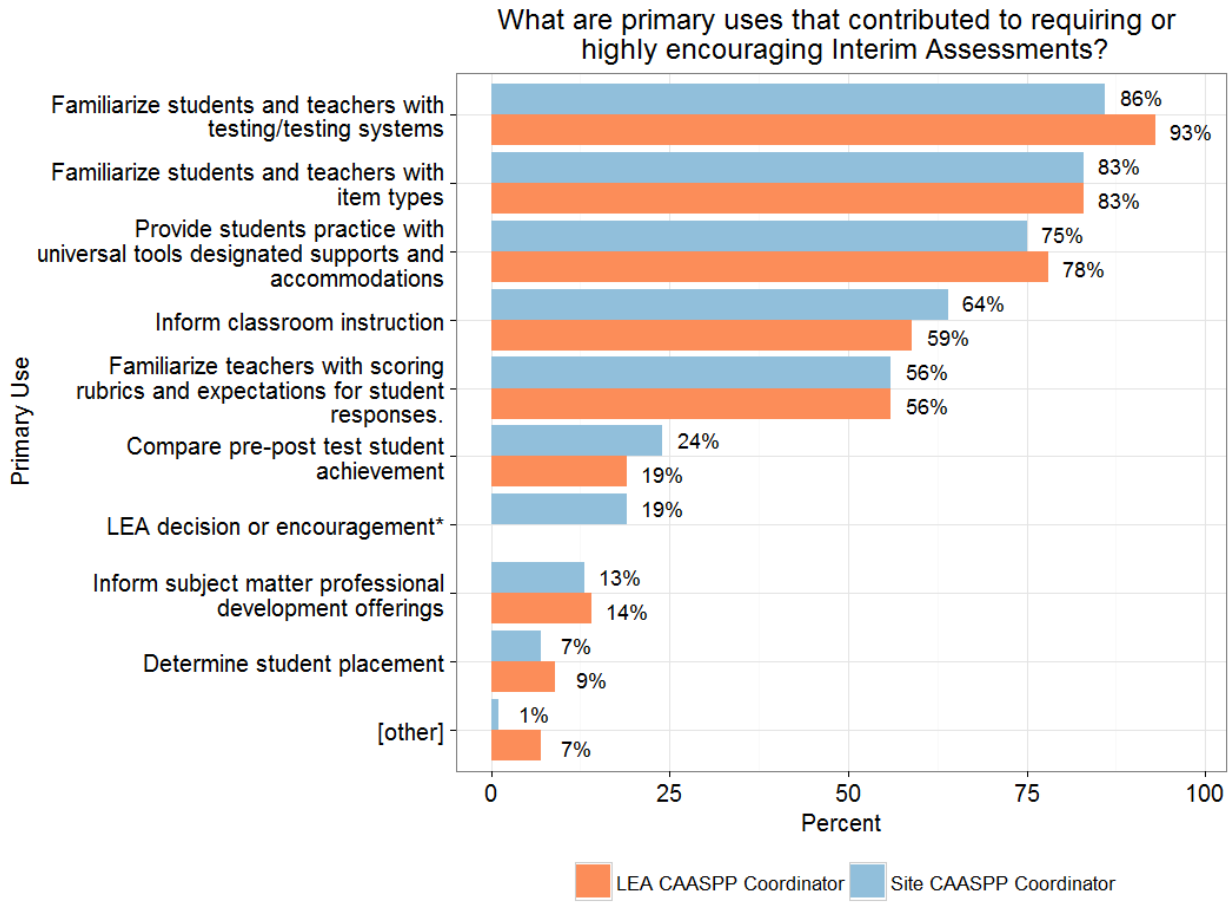


Figure Notes: * This option was not included in the LEA CAASPP Coordinator Survey.
Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.8. Primary uses that contributed to requiring or highly encouraging interim assessments (LEA and SITE surveys).

Test Administrator Uses

Because the Smarter Balanced Interim Assessments are intended to help inform instruction, we sought to know in what ways teachers administered the assessments; that is, did teachers administer the interim assessments to all students in a class or did they use these assessments to target specific subgroups of students? Overwhelmingly, test administrators reported administering the interim assessments in whole-class settings (87%); very few reported administering them only to subsets of students (Table 3.18).

Table 3.18. Interim Assessment Administration Settings Reported by TAs

TA_Q5. How did you administer the Smarter Balanced Interim Assessments to your class? Select all that apply.		
	Frequency	Percent
Whole-class (all students were given the Smarter Balanced Interim Assessments)	675	87.0%
A combination of some whole-class and some subsets of students	57	7.3%
Subset(s) only students with disabilities were assessed	50	6.4%
[other]	26	3.4%
Subset(s) only remedial students were assessed	6	0.8%
Subset(s) only English learners were assessed	3	0.4%
Subset(s) only advanced students were assessed	0	0.0%
Frequency Missing = 82.		

Because hand scoring was cited as an obstacle to using the Smarter Balanced Interim Assessments, we asked test administrators to what degree hand scoring impacted their use of the interim assessments. Of the test administrators who administered at least one assessment, hand scoring did not appear to be a factor in deciding which interim assessment to administer (75.4%). Only 16.2 percent of test administrators reported that they specifically administered interim assessments that *did not* require hand scoring.

Third-Party Assessments

We learned from the focus groups that many LEAs (a) did not use the Smarter Balanced Interim Assessments because they already used an existing third-party interim assessment platform or (b) they used the Smarter Balanced Interim Assessments, but they had some concern over the lack of integration of third-party assessment systems with the Smarter Balanced Interim Assessments. We asked LEA and school site coordinators which third-party assessment platforms they used to get a sense of usage and prevalence across the state. Just under a third (29%) of LEA coordinators and slightly less than half (46%) of school site coordinators reported not using any third-party assessment. Based on these results, there is no indication a specific third-party assessment is the preferred platform in the state (see Figure 3.9).

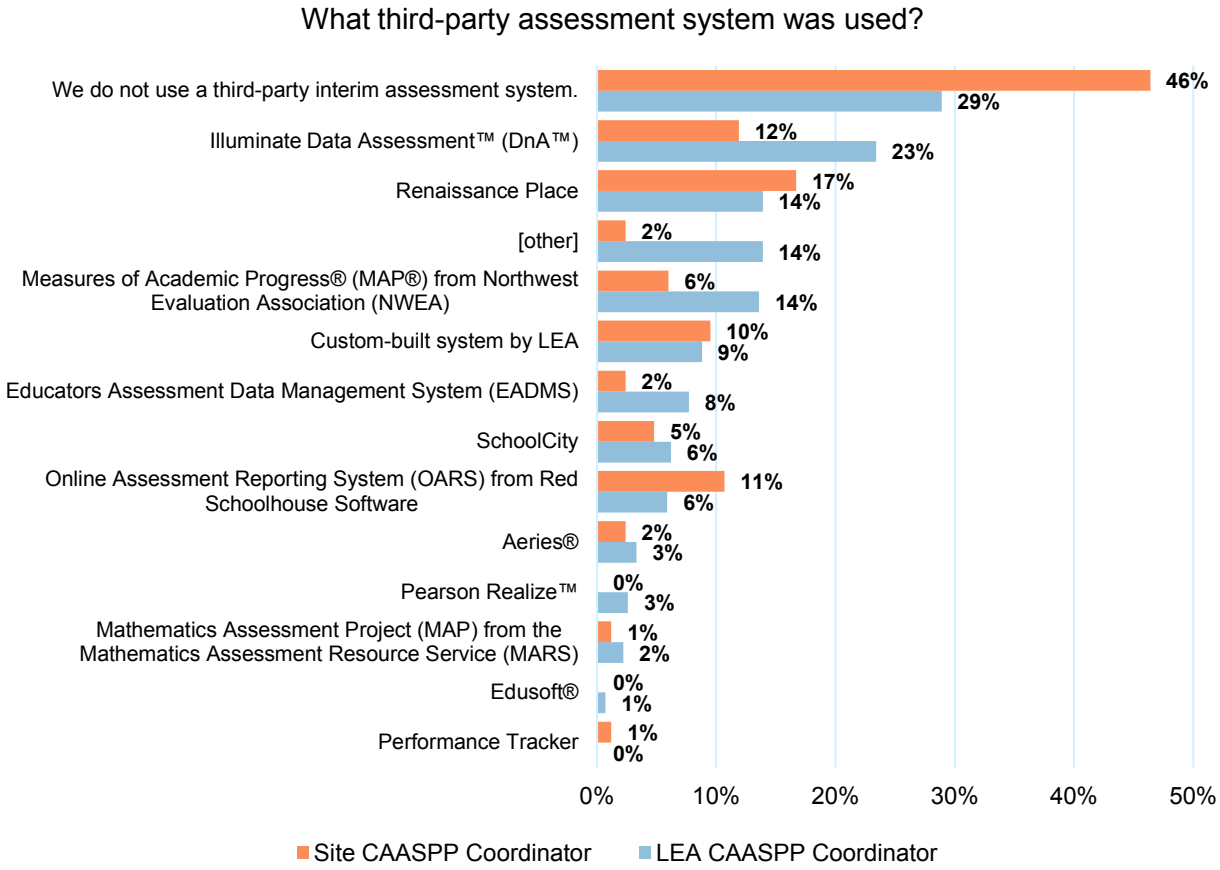


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.9. Third-party assessment usage (LEA and SITE surveys).

Preparation and Training

Table 3.19 shows the survey question numbers related to interim assessment preparation and training with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.19. Interim Assessment Preparation and Training Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q12 – Q15	34 – 37
SITE	Q15 – Q19	38 – 42
TA	Q13 – Q17	43 – 47

LEA Survey

LEA coordinators reported providing additional resources beyond those provided by CDE and Smarter Balanced. LEA-specific in-person training and/or online/print LEA-specific resources (60.4% and 42.2%, respectively) were the most commonly cited resources; only 20.7 percent of LEA coordinators reported not providing any additional resources. One-third (33.0%) of LEA coordinators reported providing some kind of hand scoring support (e.g., professional development units, stipends, and substitutes).

In terms of specific preparation and training opportunities and resources, LEA coordinators perceived the various resources as generally helpful (see Figure 3.10). Although no single resource emerged as "most helpful," the *caaspp.org* resource links and the LEA-specific training resources were reported as helpful by many LEA coordinators (63% and 60%, respectively). Approximately 66 percent of LEA coordinators found the Fall CAASPP Institute (Part 1 of the two-part training) at least somewhat helpful, while 48 percent found the Spring CAASPP Institute (Part 2 of the two-part training) at least somewhat helpful. For the "not applicable" response, it is unknown whether the resource was not consulted or did not relate to the respondent's role.

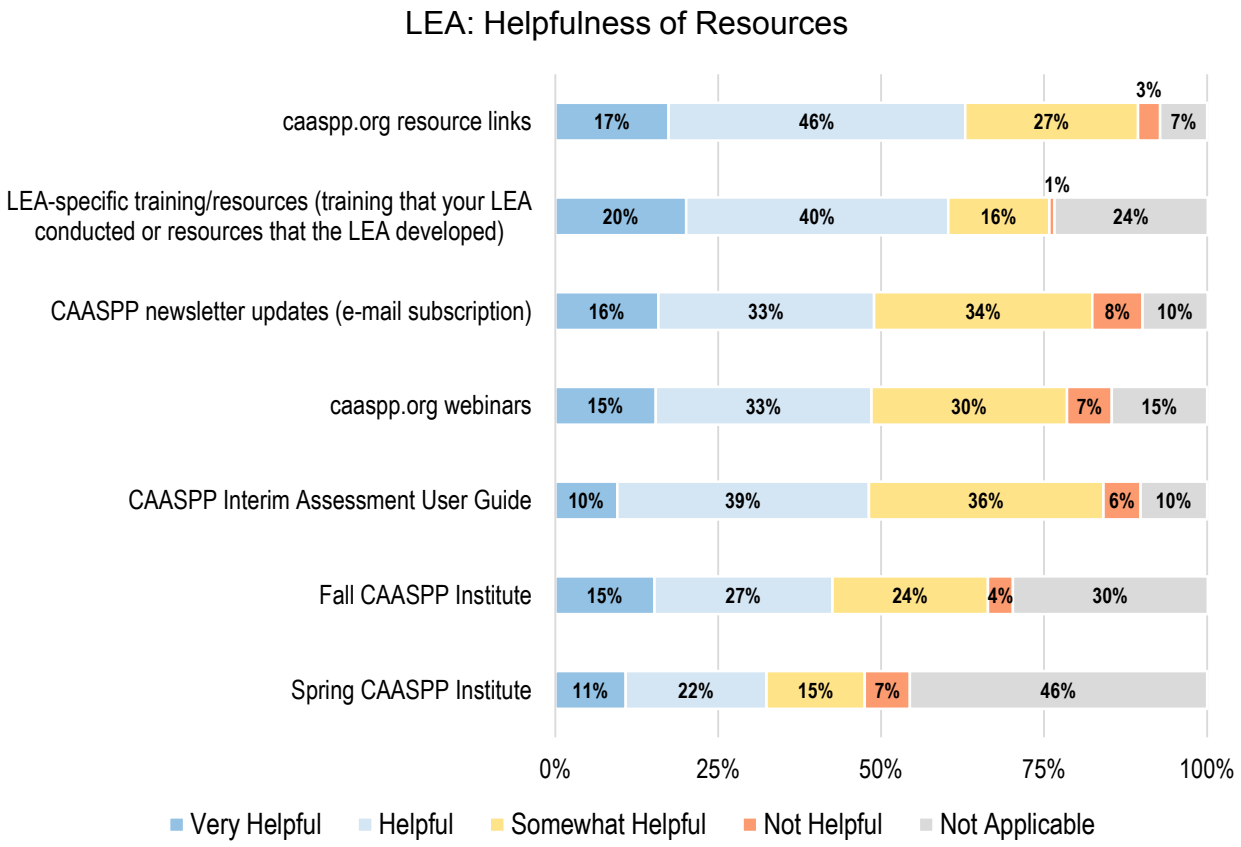


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.10. Helpfulness of Smarter Balanced Interim Assessment resources (LEA survey).

LEA coordinators reported using the various interim assessment-related systems was generally easy (see Figure 3.11). The easiest systems to use were the Test Administrator Interface (73% indicated “somewhat easy” or “very easy”), the Interim Assessment Viewing System (56%), and the Online Reporting System (ORS) (55%).

The most difficult systems to use were the Interim Assessment Reporting System (40% indicated “somewhat difficult” or “very difficult”) and Interim Assessment Hand Scoring System (37%). It is important to note that “difficult” does not necessarily imply fault within the system. We did not survey stakeholder knowledge of how each component is designed to work. These difficulties are relative to how the users interpreted their experience, regardless if it aligned with the intended use. Of the coordinators who indicated the Interim Assessment Reporting System was “very difficult” to use, they most commonly cited issues with the system, the level of report detail, and downloading results as reasons for the difficulty (n=28). Most of their open-ended comments related to the level of detail, such as the reports not being detailed enough to inform instruction (n=10) and downloading results (n=12). Others suggested that filtering results by teacher, not just grade (n=6), was not possible nor was downloading district-wide data (n=6). Although coordinators in our sample translated these shortcomings as the system being “difficult” to use, Smarter Balanced intentionally omitted reporting a level of detail that could not be supported psychometrically and that might lead to invalid inferences of student learning. Longer tests would be required to make valid inferences at a more detailed level.

Among those who indicated the Interim Assessment Hand Scoring System was “very difficult” to use, the system being cumbersome was the most commonly cited reason (n=12). Users found the process of scoring to be difficult, time consuming, and not user-friendly. Similarly, five additional comments were related to system inefficiency—they were not able to sort items in the desired order and they were unable to sort by site. Only two respondents noted dissatisfaction with the interface specifically, and the inability to report at the school or district level. Only one respondent noted the system improved from the previous (2014–15) school year.

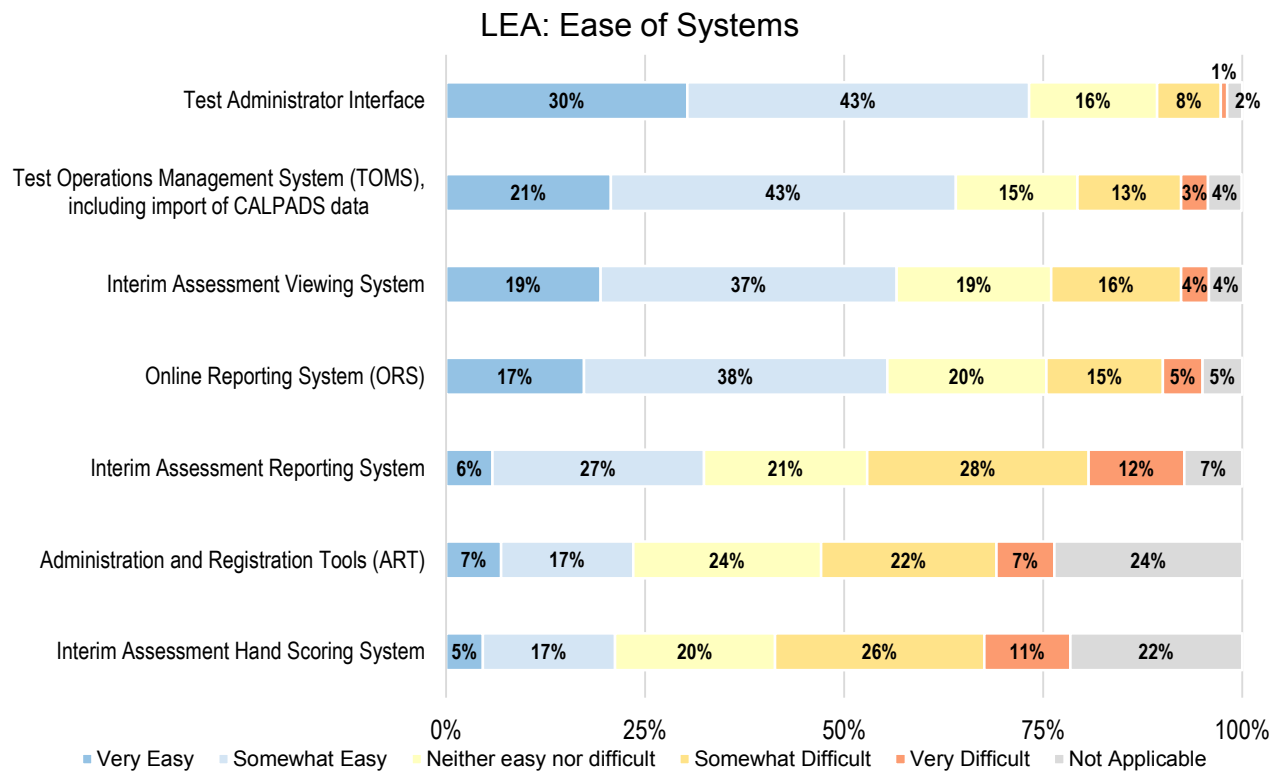


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.11. Ease of Smarter Balanced Interim Assessment Systems (LEA survey).

SITE Survey

A majority of school site coordinators reported that their LEA and/or school provided some kind of additional resources beyond what CDE and Smarter Balanced provided (73.4%). Approximately half of the school site coordinators reported being offered in-person training (46.8%). In-person training only (19.0%) and in-person training plus online and print resources (15.2%) were the most commonly cited opportunity combinations.

In terms of specific preparation and training opportunities and resources, school site coordinators generally perceived the various resources as helpful (see Figure 3.12). Although no single resource emerged as being the "most helpful," 62 percent of school site coordinators reported the LEA training and resources to be most helpful; 68 percent of them cited the school-specific training and resources as being the most helpful. Slightly less than half of them reported the Fall and Spring CAASPP Institutes (42% and 41%, respectively) as at least somewhat helpful, and a similar percentage (47%) found the caasp.org Webinars at least somewhat helpful. Additionally, almost half of the school site coordinators reported reading/receiving the CAASPP newsletter updates (48%).

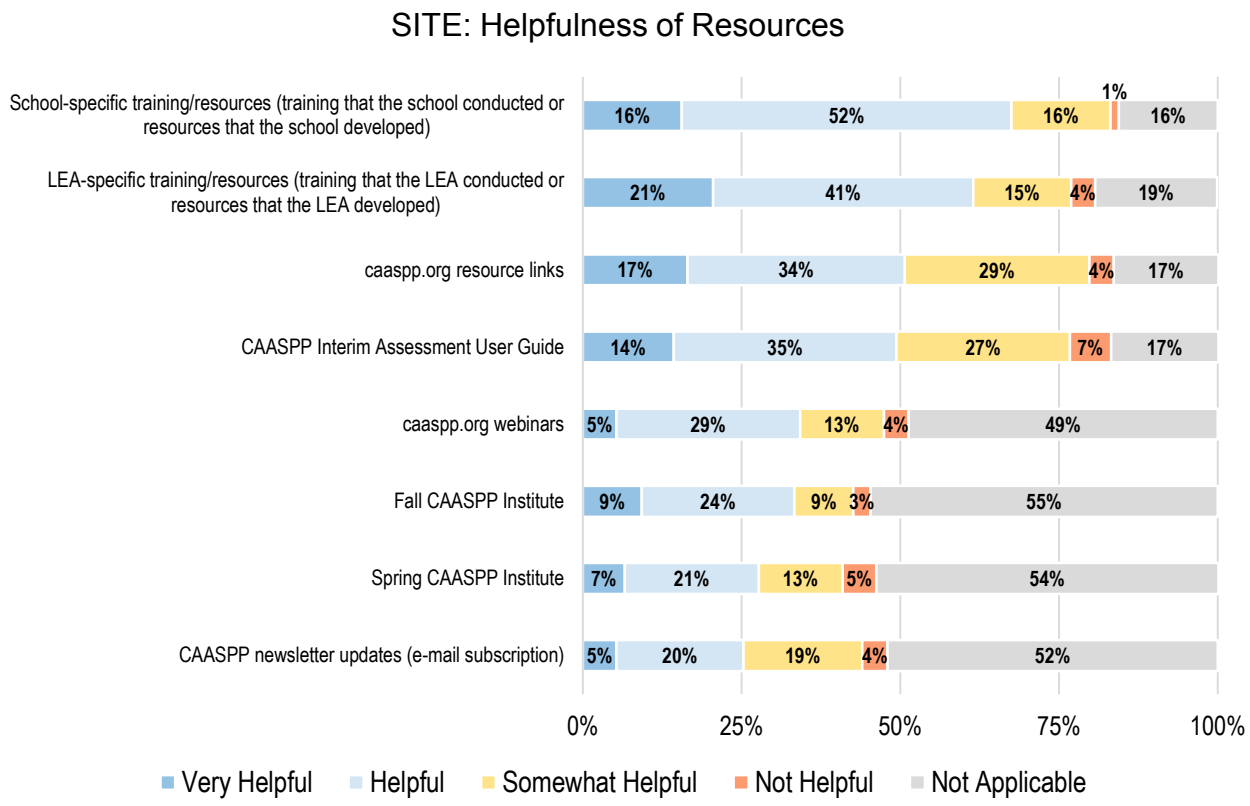


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.12. Helpfulness of Smarter Balanced Interim Assessment Resources (SITE survey).

School site coordinators generally reported using the various interim assessment-related systems was easy. The easiest system to use was the Test Administrator Interface (83.5% indicated “somewhat easy” or “very easy”). The most difficult system to use was the Interim Assessment Hand Scoring System (30.4%), indicated “somewhat difficult” or “very difficult”); however, a similar percentage also reported not using it, as indicated by their “not applicable” rating (30.4%). Among those who indicated the system was “very difficult” to use, the most commonly cited reasons for the difficulty were that the system was confusing to navigate and had unclear directions.

To follow up on some of the challenges in using the Interim Assessment Reporting System expressed during the focus groups, we asked site coordinators how easy it was to perform certain functions within the system (see Figure 3.13). No clear patterns emerged; school site coordinators did not report certain functions were particularly easy or more difficult than others. Some functions were used less than others (as indicated by the reported percentage of “not applicable”). For example, many school site coordinators reported viewing student results by grade (74.4%); fewer reported comparing results to prior years (45.0%).

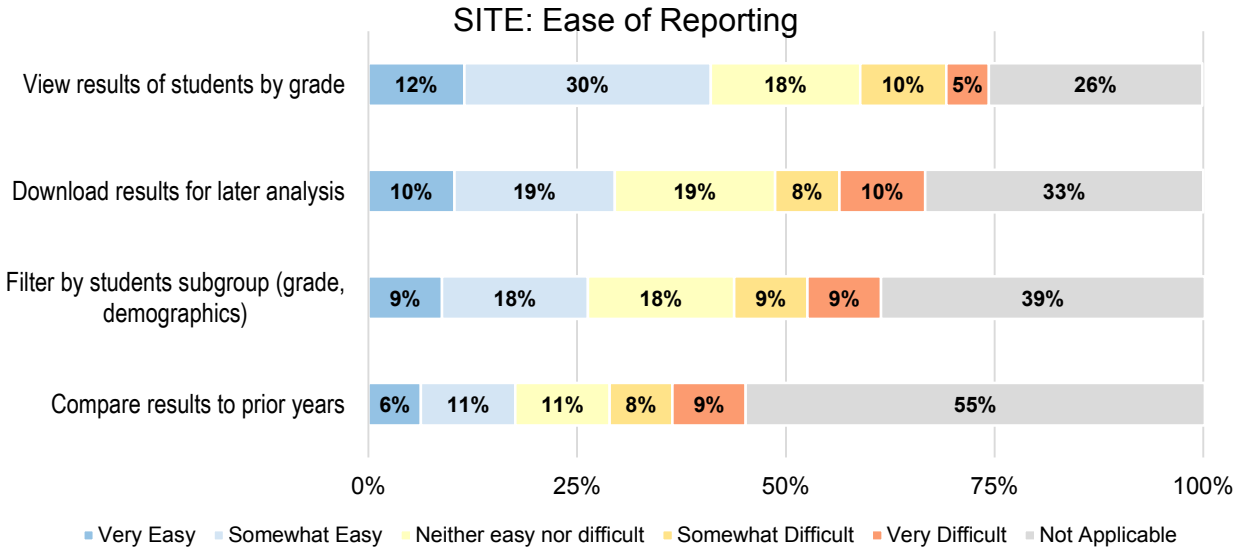


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.13. Ease of Smarter Balanced Interim Assessment Reporting System (SITE survey).

TA Survey

Many test administrators reported their school or LEA provided some kind of additional resources beyond those provided by CDE and Smarter Balanced (62.9%); however, a substantial percentage of the test administrators was not sure what was provided (37.1%) and only a few (8.9%) stated they were provided support for hand scoring (e.g., professional development units, stipends, substitutes).

As indicated by the high percentages of "not applicable" ratings, many test administrators did not use many of the available resources (see Figure 3.14). For example, 81.1 percent of test administrators did not attend the Fall CAASPP Institutes and 78.0 percent did not attend the Spring Institute. They reported the most used resources were the CAASPP Interim Assessment User Guide (70.7%), school-specific training/resources (78.8%), and caaspp.org resource links (66.2%). They generally found these resources to be helpful.

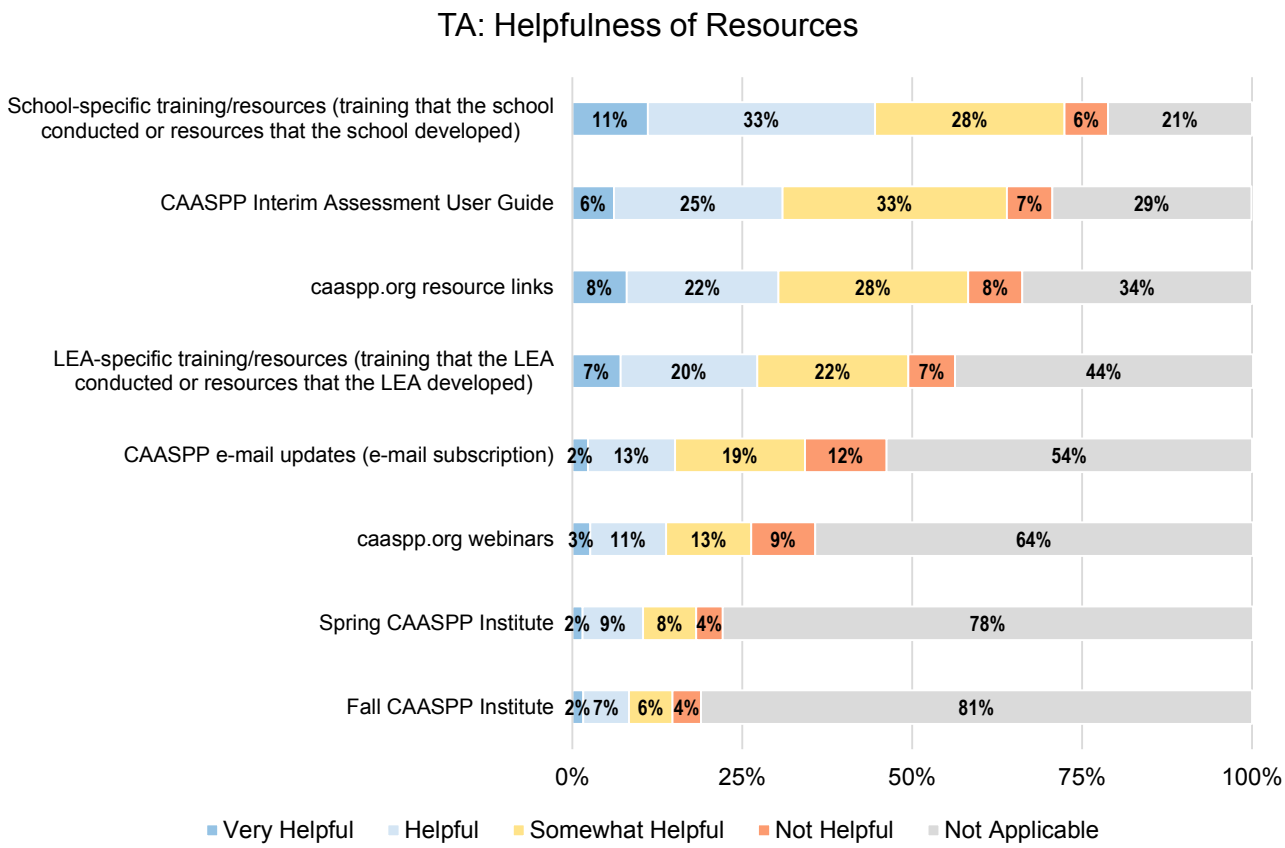


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.14. Helpfulness of Smarter Balanced Interim Assessment resources (TA survey).

Test administrators generally reported the various systems were easy to use. The easiest system to use was the Test Administrator Interface (64.9% indicated “somewhat easy” or “very easy”). No systems were reported to be particularly more difficult to use than the others (Figure 3.15).

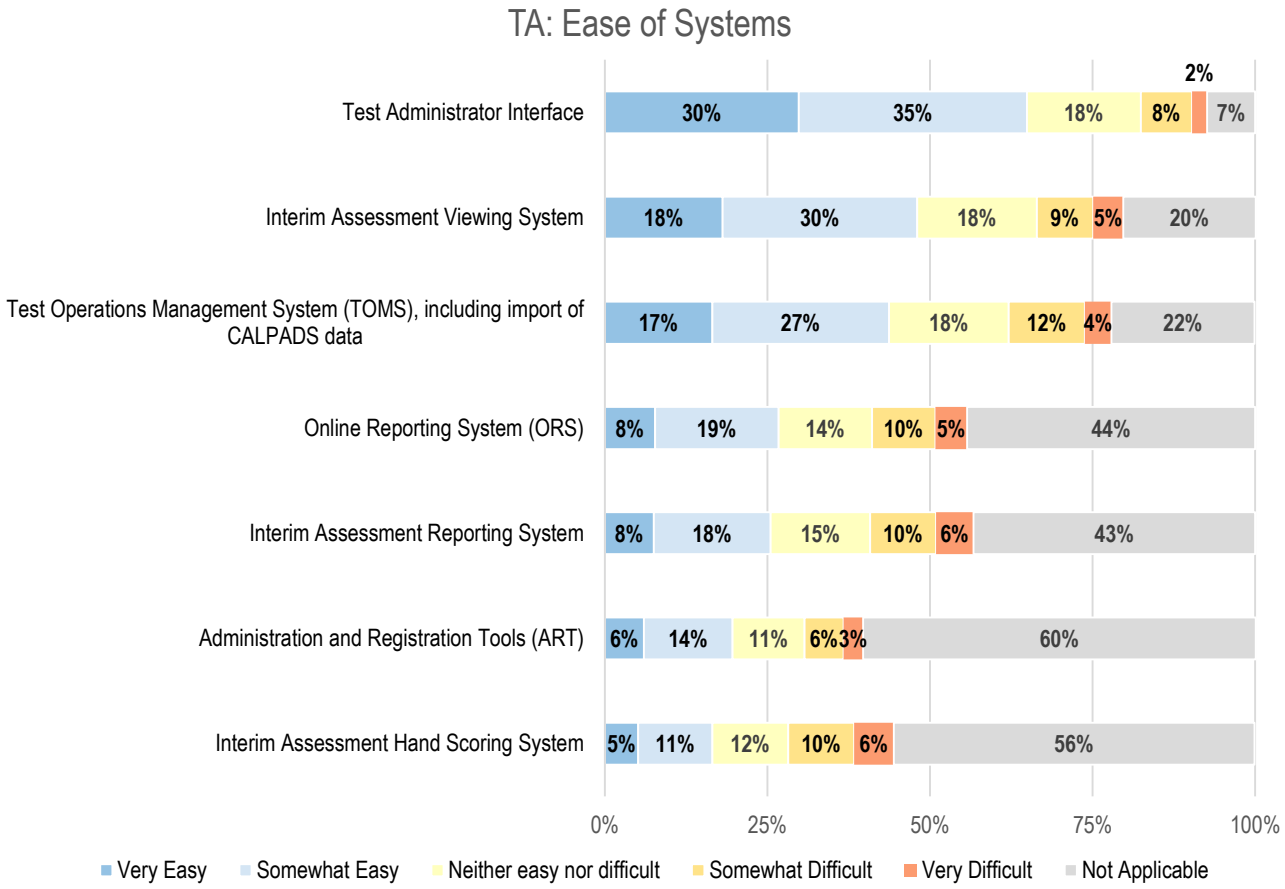


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.15. Ease of Smarter Balanced Interim Assessment Systems (TA survey).

As with the school site coordinators, we asked test administrators how easy it was to perform certain functions within the Interim Assessment Reporting System (Figure 3.16). No clear patterns emerged in terms of relative ease; however, approximately half of the test administrators reported not using most of the functions (as indicated by the "not applicable" rating). Because LEA coordinators control the functions and permissions to the Interim Assessment Reporting System via the Administration and Registration Tools (ART) system, the high percentage of TAs that reported not using most of the functions may reflect lack of permission, rather than lack of use.

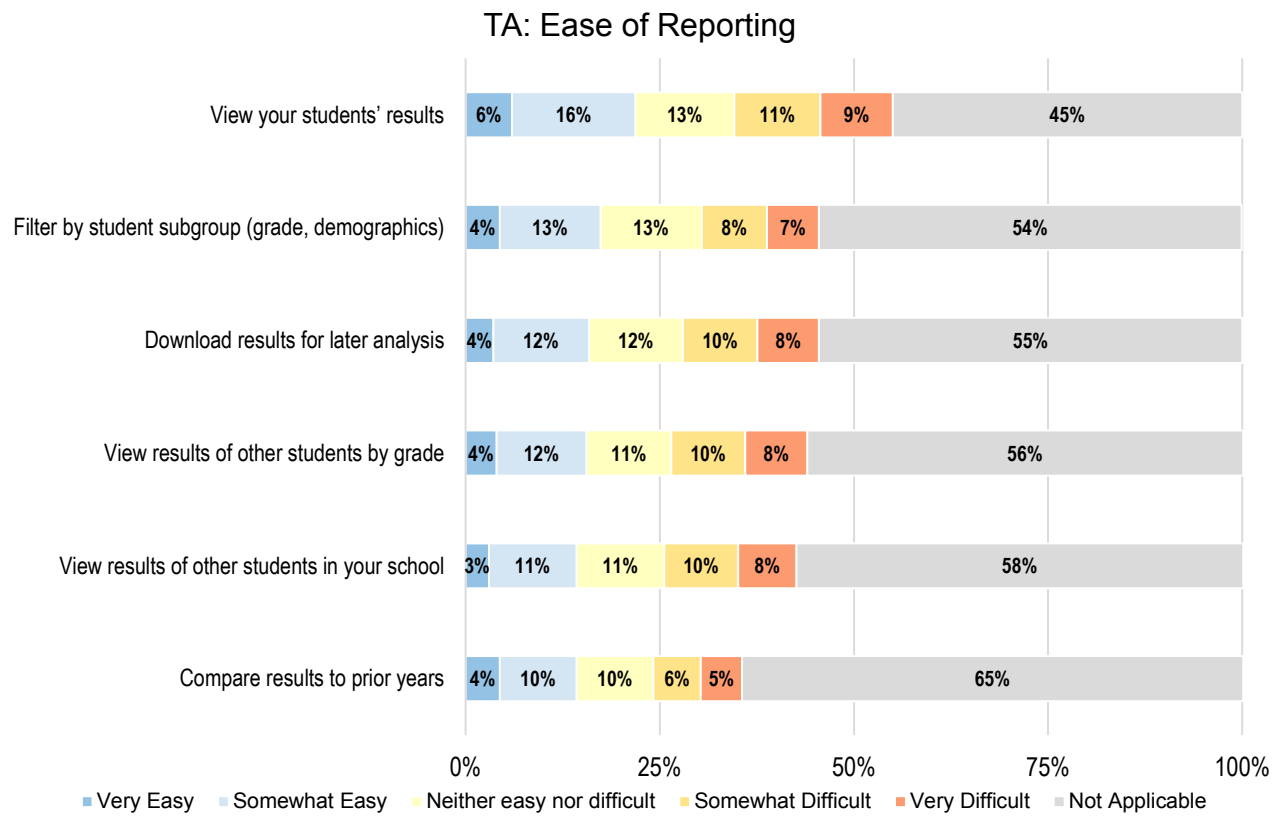


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.16. Ease of Smarter Balanced Interim Assessment Reporting System (TA survey).

Impacts on Instruction

Table 3.20 shows the survey question numbers related to interim assessment preparation and training with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.20. Interim Assessment Impacts on Instruction Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q8 – Q11	48 – 50, 56
SITE	Q11 – Q14	51 – 53, 56
TA	Q10 - Q12	54 – 55, 58

LEA Survey

LEA coordinators reported that familiarizing students and teachers with testing/testing systems was the most useful benefit of administering the Smarter Balanced Interim Assessments (84% indicated “very useful” or “useful”), followed by familiarizing teachers with scoring rubrics and expectations for student responses (61%). They indicated the least useful activities were determining student course placement (47% indicated “not useful”) and promoting differentiated instruction (38%) (see Figure 3.17).

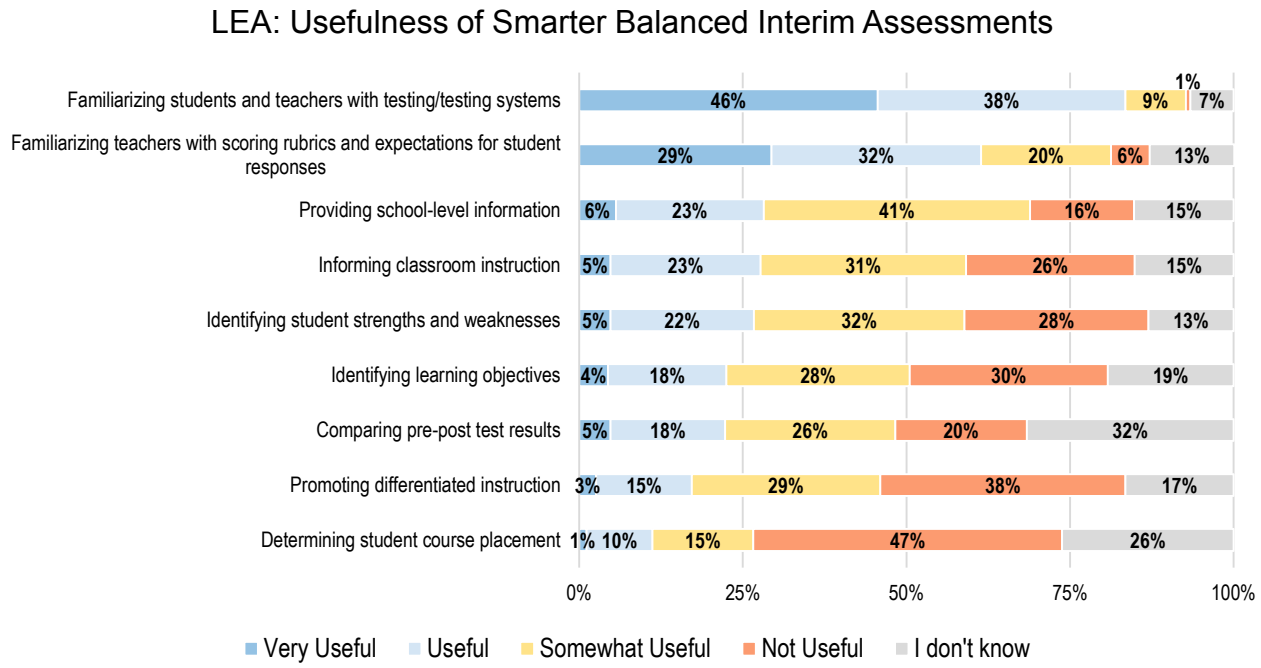


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.17. Usefulness of Smarter Balanced Interim Assessments (LEA survey).

Approximately one-half of the LEA coordinators stated their decisions about students and LEA-level educator support topics were not based on the Smarter Balanced Interim Assessments (52% and 49%, respectively). However, for the LEA coordinators that did report results-based decisions, approximately one-third noted identifying classroom or grade-level strengths and weaknesses (33%) and identifying student strengths and weaknesses (32%) as being decided on the basis of results (see Figure 3.18).

The CDE and Smarter Balanced recommend that Smarter Balanced Interim Assessment results be used in combination with other indicators about student performance. It is unclear whether respondents interpreted the phrasing of the survey questions (e.g., “What types of decisions about students will be or have been made based on the Smarter Balanced Interim Assessment Results”) as referring to decisions made solely using Smarter Balanced Interim Assessment results or decisions that used the results along with other data points.

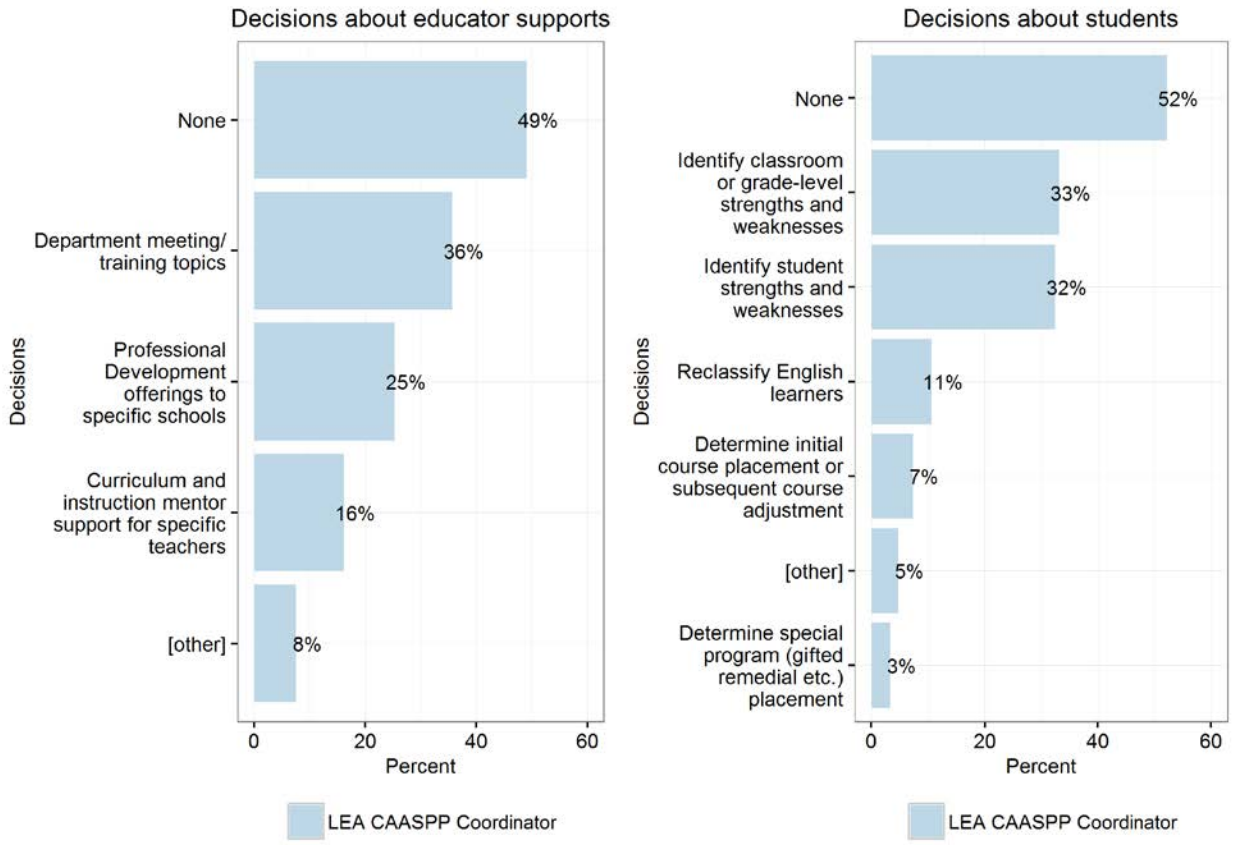


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.18. Decisions about educator supports and students (LEA survey).

SITE Survey

School site CAASPP coordinators reported that familiarizing students and teachers with testing/testing systems was the most useful benefit of administering the Smarter Balanced Interim Assessments (87% indicated “very useful” or “useful”). The second most useful result of the testing was familiarizing teachers with scoring rubrics and expectations for student responses (60%). The least useful activities were determining student course placement (59% indicated “not useful”) and promoting differentiated instruction (40%) (Figure 3.19).

Slightly less than half of the school site coordinators stated their decisions about students and school-level educator support topics were not being made based on the Smarter Balanced Interim Assessments (44% and 47%, respectively). For the school site coordinators that did report results-based decisions, identifying classroom or grade-level strengths and weaknesses was more commonly reported than was identifying student strengths and weaknesses (Figure 3.20).

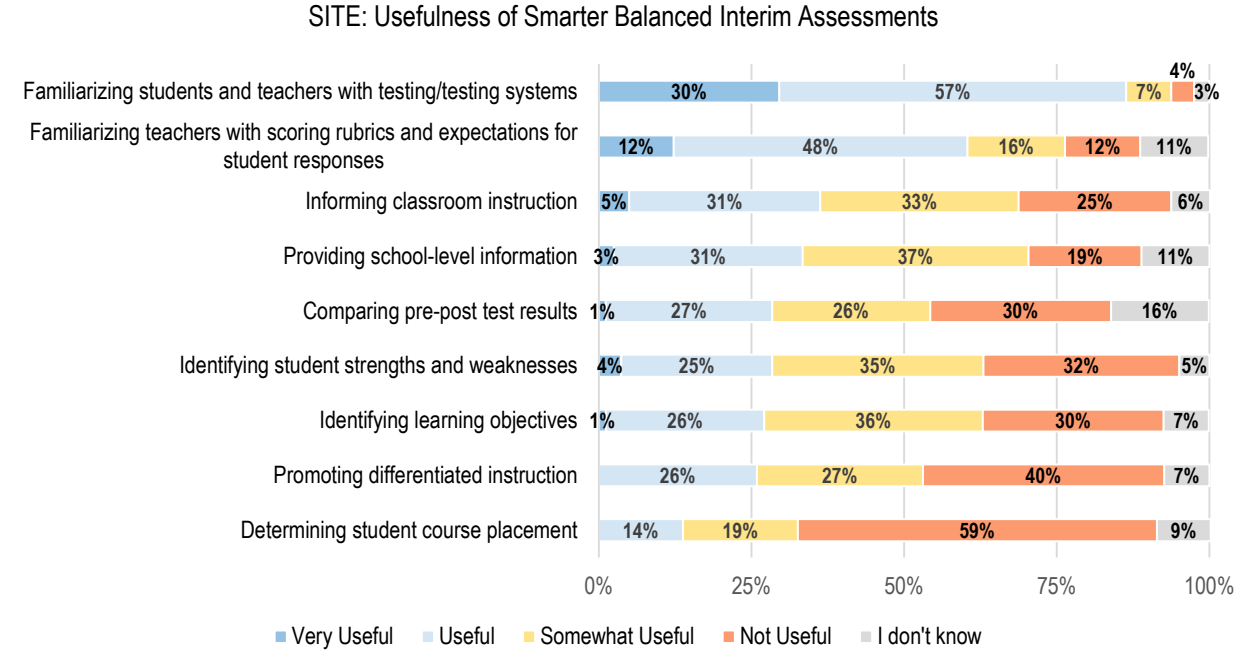


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.19. Usefulness of Smarter Balanced Interim Assessments (SITE survey).

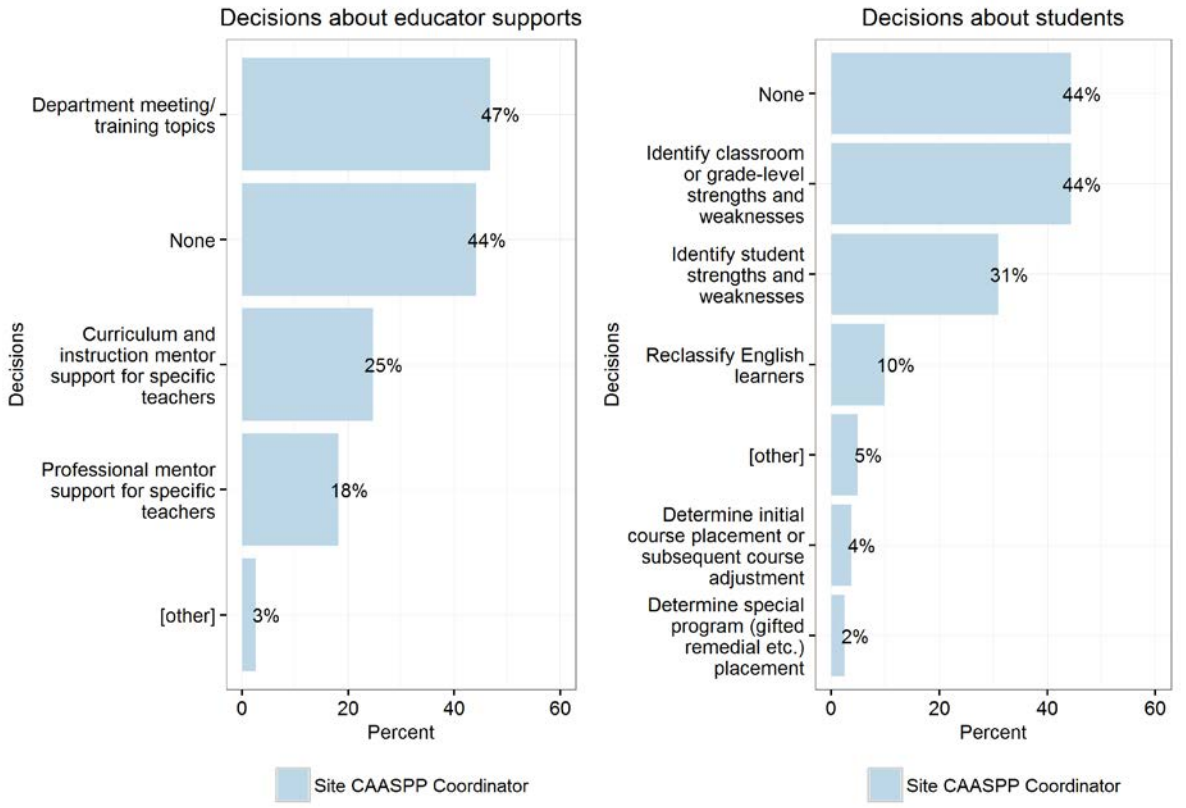


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.20. Decisions about students and educator support (SITE survey).

TA Survey

Test administrators reported that familiarizing students and teachers with testing/testing systems was the most useful benefit of administering the Smarter Balanced Interim Assessments (62.8% indicated "very useful" or "useful"), followed by familiarizing students and teachers with item types (56.4%). The least useful activities were determining student course placement (39.8% indicated "not useful") and promoting differentiated instruction (35.1%) (Figure 3.21).

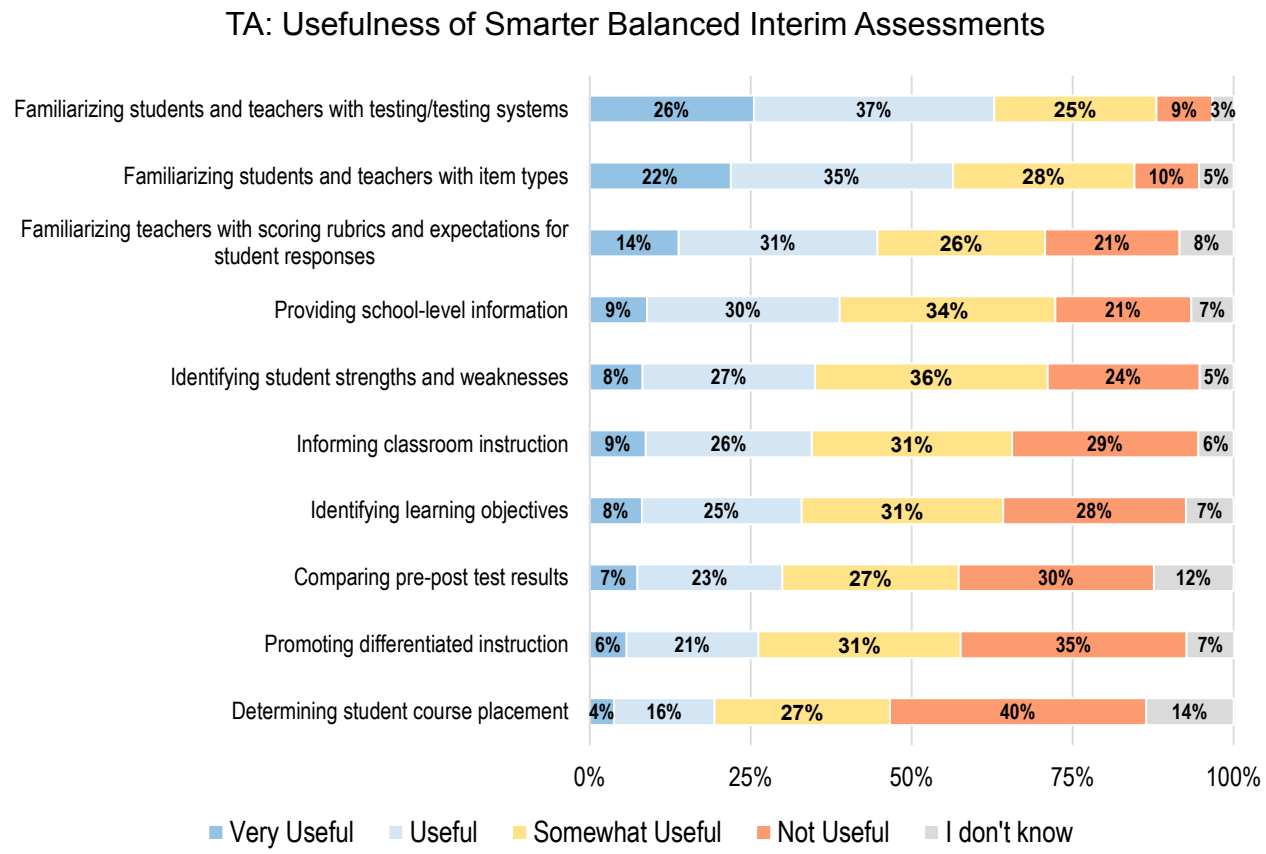


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.21. Usefulness of Smarter Balanced Interim Assessments (TA survey).

As shown in Figure 3.22, identifying classroom or grade-level strengths and weaknesses (54.1%) and identifying student strengths and weaknesses (53.5%) were most often reported by test administrators as the types of decisions that teachers made about students based on the Smarter Balanced Interim Assessments.

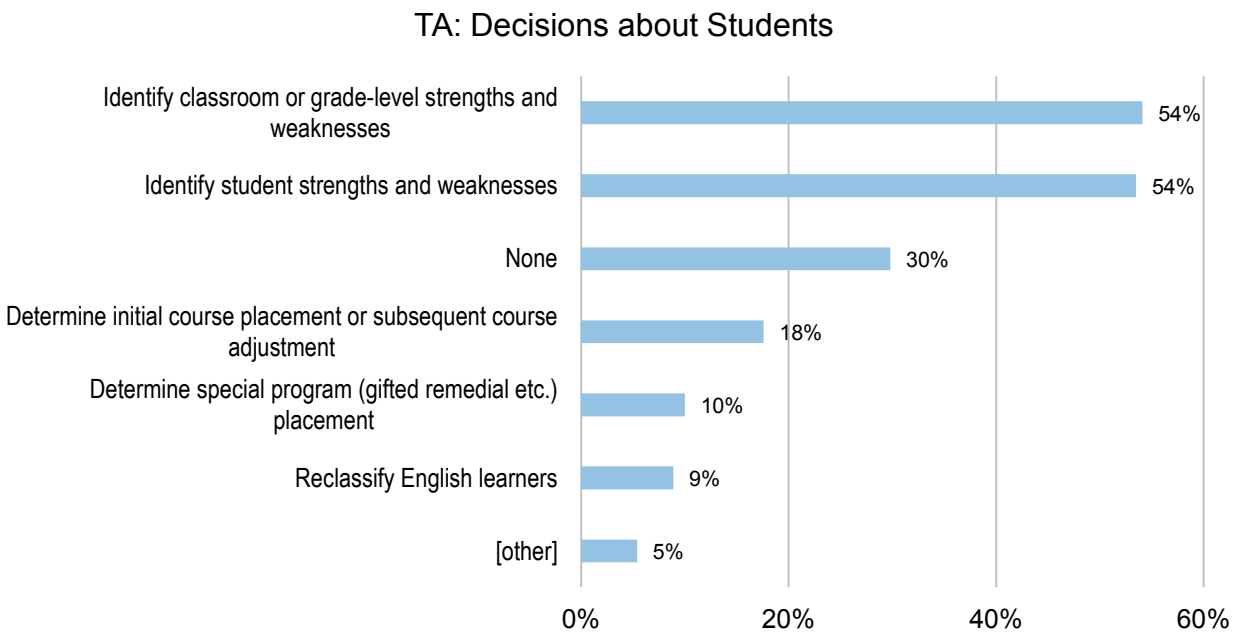


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.22. Decisions about students (TA survey).

Confidence in Results

LEA and school site coordinators reported having some or a great deal of confidence the Smarter Balanced Interim Assessments provide valid indicators of student achievement (74.0% and 70.9%, respectively). Although approximately half of test administrators reported having some or a great deal of confidence that the results are valid indicators of student achievement (56.0%), 44.0% of them reported having very little to no confidence in the validity of the results (see Figure 3.23).

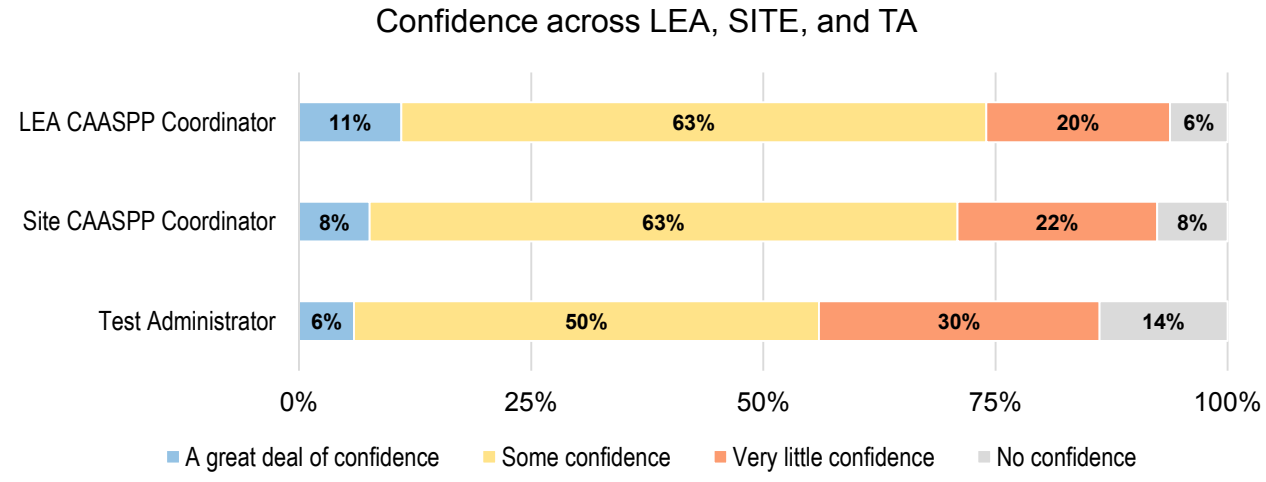


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.23. Confidence in the validity of results (LEA, SITE, and TA surveys).

Assessment Administration

Because LEA coordinators are more removed from actual administration of the Interim Assessments, questions about administration were provided only to the site coordinators and test administrators.

Table 3.21 shows the survey question numbers related to interim assessment preparation and training with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.21. Interim Assessment Administration Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	NA	NA
SITE	Q8	59
TA	Q6-Q7	60 – 61

School site coordinators were mixed in their responses regarding the extent to which test administration procedures outlined in the interim assessment resources were followed. Approximately one-third of the school site coordinators indicated each of “extreme”, “moderate”, and “slight” degrees to which they perceived test administration procedures were followed as outlined in the CAASPP Interim Assessment User Guide (Table 3.22).

Table 3.22. Reported Standardization of Administration Procedures by School Site Coordinators

SITE_Q8. To what degree did your school attempt to standardize how the Smarter Balanced Interim Assessments were administered?		
	Frequency	**Percent**
Extreme: We required formal procedural training and monitored compliance with the procedures outlined in the CAASPP Interim Assessment User Guide.	32	37.2%
Moderate: Some but not all components from the procedures outlined in the CAASPP Interim Assessment User Guide were covered in formal training and required to be followed.	28	32.6%
Slight: Teachers were allowed great flexibility in how interim assessments were administered based on the procedures outlined in the CAASPP Interim Assessment User Guide.	24	27.9%
[other]	2	2.3%
Frequency Missing = 4		

More test administrators (51.9%) reported “extreme” attempts to follow administration procedures as outlined than reported “slight” attempts to follow them (16.1%) (Table 3.23).

Table 3.23. Reported Standardization of Administration Procedures by Test Administrators

TA_Q6. To what degree did your school attempt to standardize how the Smarter Balanced Interim Assessments were administered?	Frequency	Percent
Extreme: We required formal procedural training and monitored compliance with the procedures outlined in the CAASPP Interim Assessment User Guide.	393	51.9%
Moderate: Some but not all components from the procedures outlined in the CAASPP Interim Assessment User Guide were covered in formal training and required to be followed.	232	30.6%
Slight: Teachers were allowed great flexibility in how Smarter Balanced Interim Assessments were administered based on the procedures outlined in the CAASPP Interim Assessment User Guide.	122	16.1%
[other]	10	1.3%
Frequency Missing = 106		

We asked test administrators to report to what degree they followed the procedures outlined in the CAASPP Interim Assessment User Guide. The large majority of administrators (greater than 80%) reported always (a) allowing the full testing time, (b) ensuring that ancillary materials were used appropriately, (c) monitoring student progress, and (d) ensuring students worked independently (see Figure 3.24).

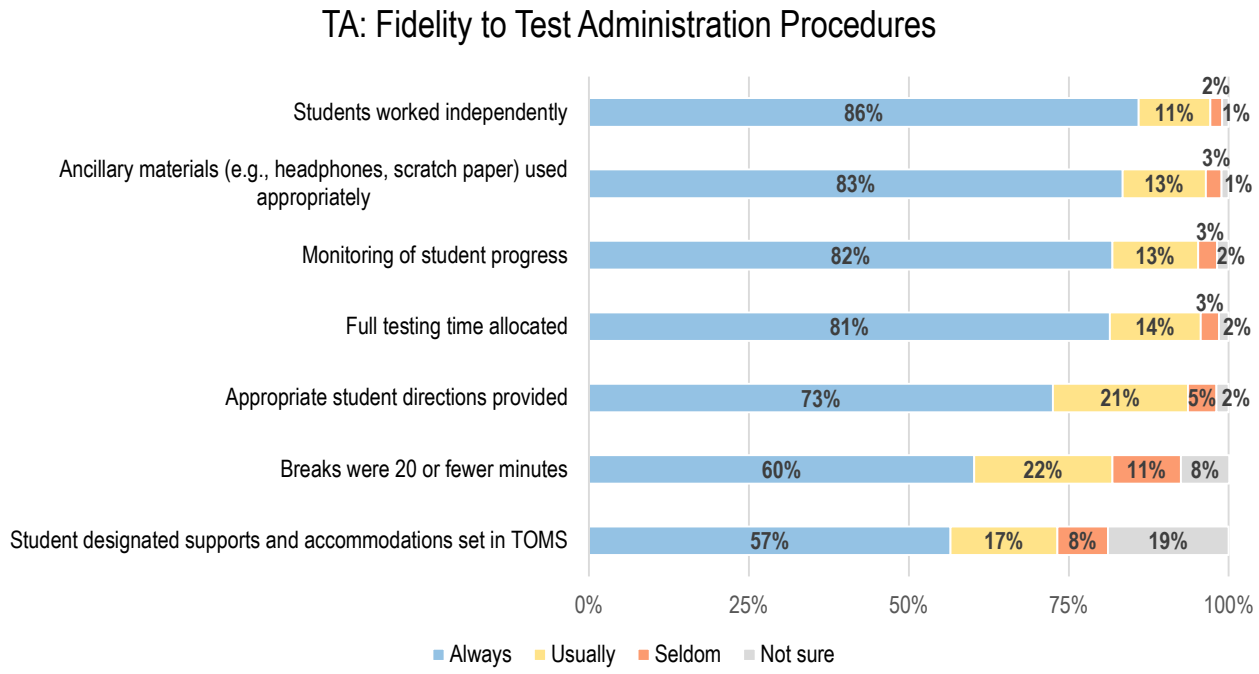


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.24. Fidelity to test administration procedures (TA survey).

Accommodations and Supports

Table 3.24 shows the survey question numbers related to accommodations and supports on the Smarter Balanced Interim Assessments with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.24. Interim Assessment Accommodations Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	NA	NA
SITE	Q9 – Q10	62 – 63
TA	Q8 – Q9	64 – 65

Because LEA coordinators are more removed from actual administration of the interim assessments, questions about accommodations and supports were provided only to the site coordinators and test administrators. The site coordinators and test administrators were asked about how well the Smarter Balanced Interim Assessments incorporate supports for students with disabilities (SWD) and English learners (ELs).

SITE Survey

Although school site coordinators generally agreed the accommodations, supports, and features of the interim assessments contributed to positive experiences for SWDs, there was not overwhelming agreement. Their highest level of agreement was that the Smarter Balanced Interim Assessments improved access to the test (66.3% strongly agreed or agreed) and that the supports functioned properly during testing (64.0%) (see Figure 3.25).

TA Survey

Test administrators generally agreed the accommodations and features of the interim assessments contributed to positive experiences for SWDs; however, not all among them agreed (see Figure 3.26).

For example, less than half of test administrators agreed that for SWDs, the assessments (a) improved access to the test, (b) are consistent with instructional practices used in the classroom, and (c) functioned properly during operational testing. Additionally, a large percentage of test administrators reported not knowing whether the Smarter Balanced Interim Assessments resulted in positive experiences for SWDs (as indicated by “I don’t know”). Test administrators also indicated uncertainty regarding whether the supports and features of the Interim Assessments contributed to positive experiences for ELs (see Figure 3.26).

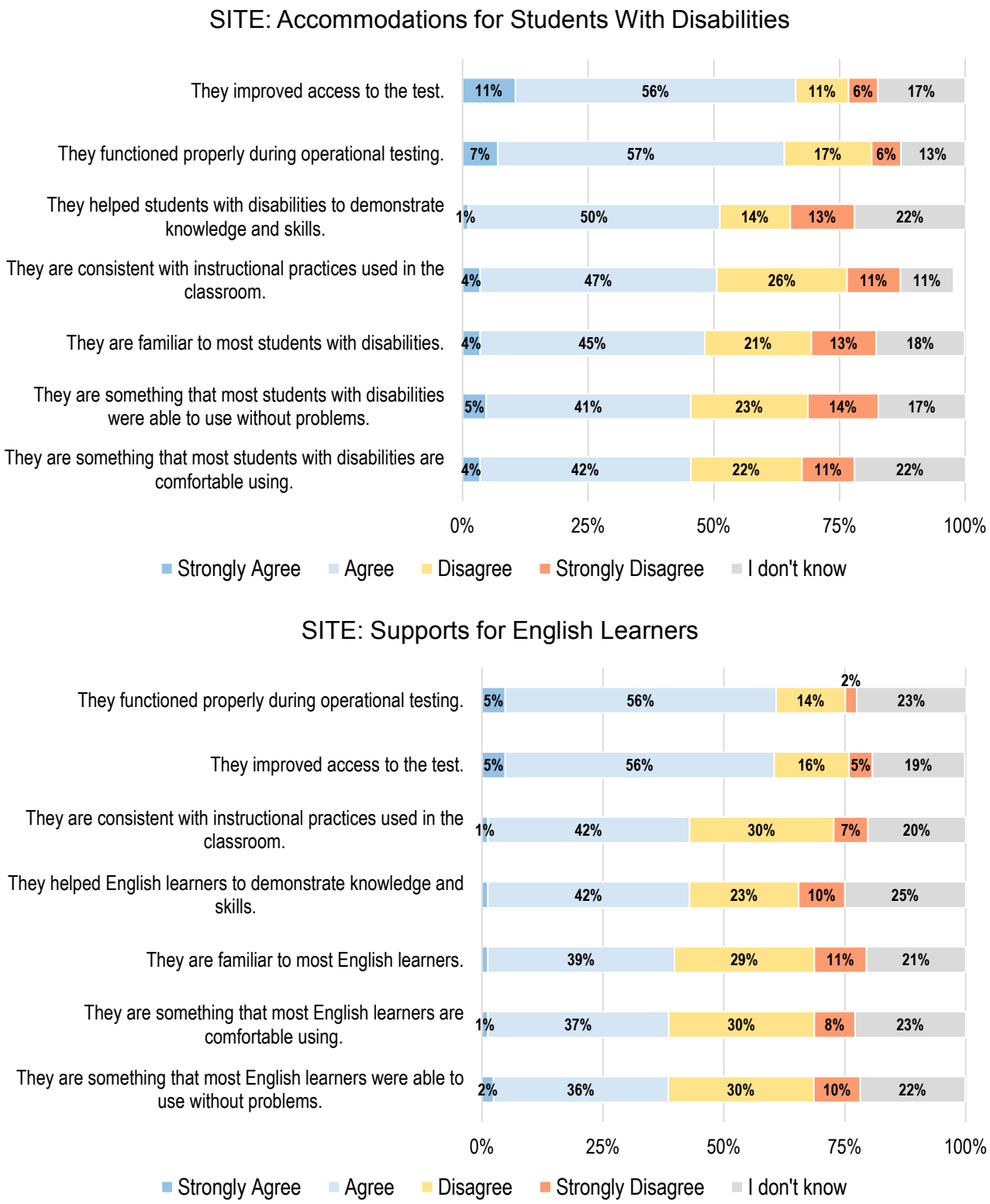


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.25. Accommodations for students with disabilities and supports for English learners (SITE survey).

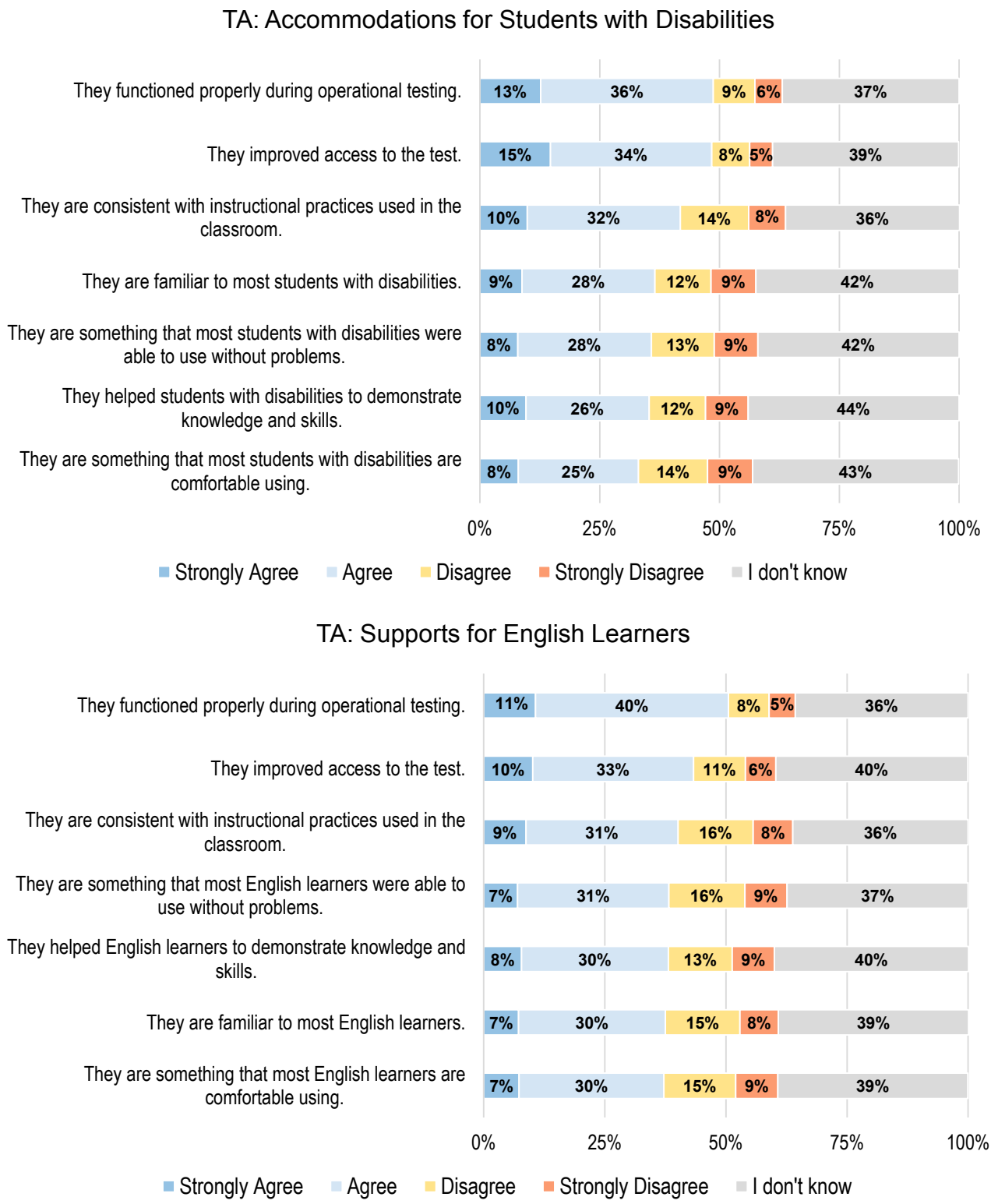


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.26. Accommodations for students with disabilities and supports for English learners (TA survey).

Challenges

Each stakeholder group indicated whether any activities presented challenges related to the topics of administration, test content, and hand scoring. Table 3.25 shows the survey question numbers related to challenges with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.25. Interim Assessment Challenges Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q16 – Q19	66 – 69
SITE	Q18 – Q23	70 – 73
TA	Q18 – Q21	74 – 77

LEA Survey

The least challenging administration activity, according to 58 percent of respondents at the LEA level, was setting up the interim assessment for administration. The most challenging activity was determining the appropriate designated supports and accommodations (28% indicated this was a major challenge) (see Figure 3.27).

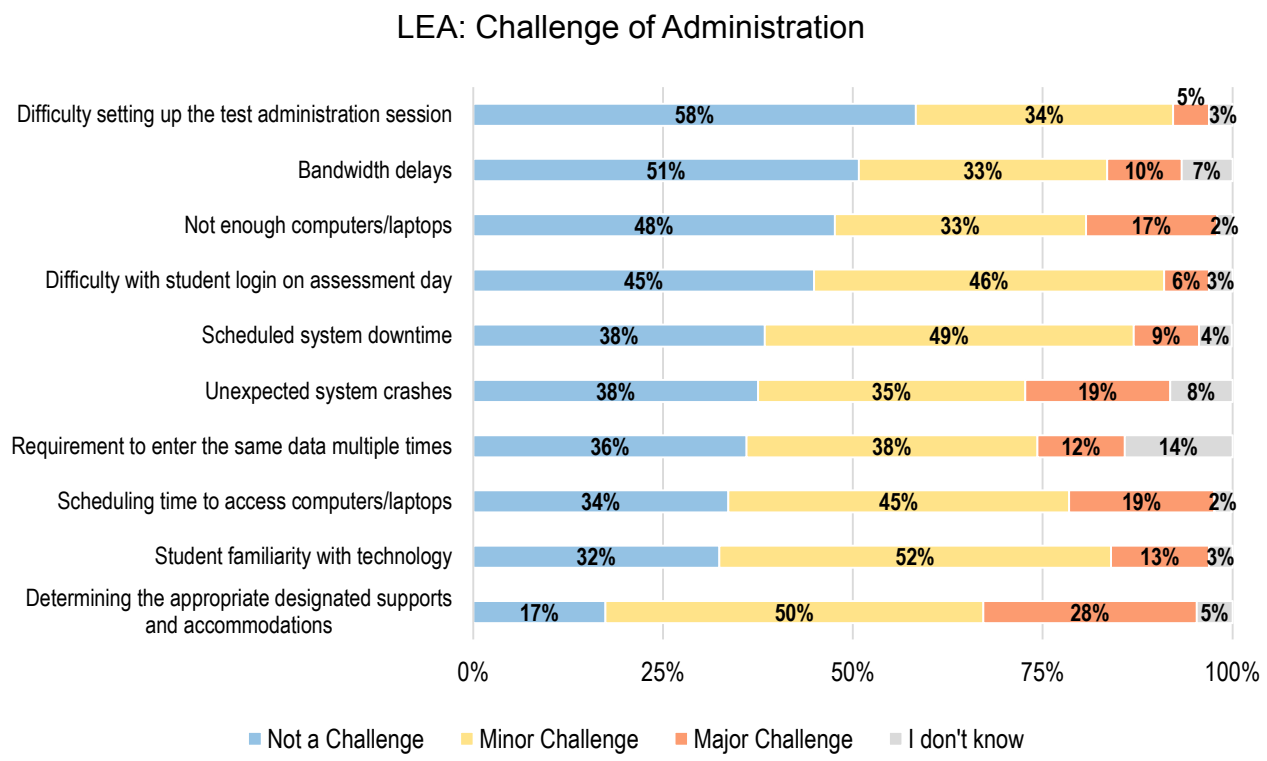


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.27. Challenge of administration (LEA survey).

Few LEA coordinators reported that the content of the assessments was a challenge (too easy or too difficult for results to be helpful, did not match curriculum, or confusing item types) (see Figure 3.28).

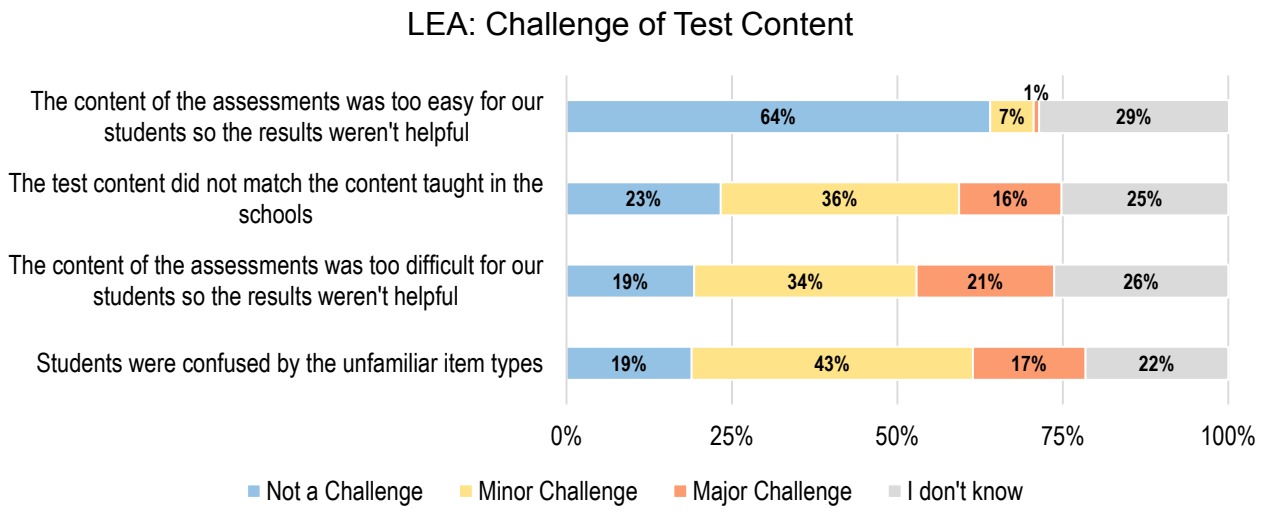


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.28. Challenge of test content (LEA survey).

Although the hand scoring activities of Smarter Balanced Interim Assessments can facilitate professional learning for educators and help them better understand the system of scoring—including awareness of rubrics that can be used in the classroom for teaching and student learning— some respondents in our sample cited the hand scoring aspect of interim assessments as a challenge. LEA coordinators reported that committing adequate time to complete scoring was the greatest challenge (48%). Identifying classroom teachers willing to hand score their own students' responses was also noted as a major challenge (33%). Many LEA coordinators reported not knowing whether some of the hand scoring activities were challenging or not, suggesting they either did not participate in these activities and/or did not conduct any hand scoring (see Figure 3.29).

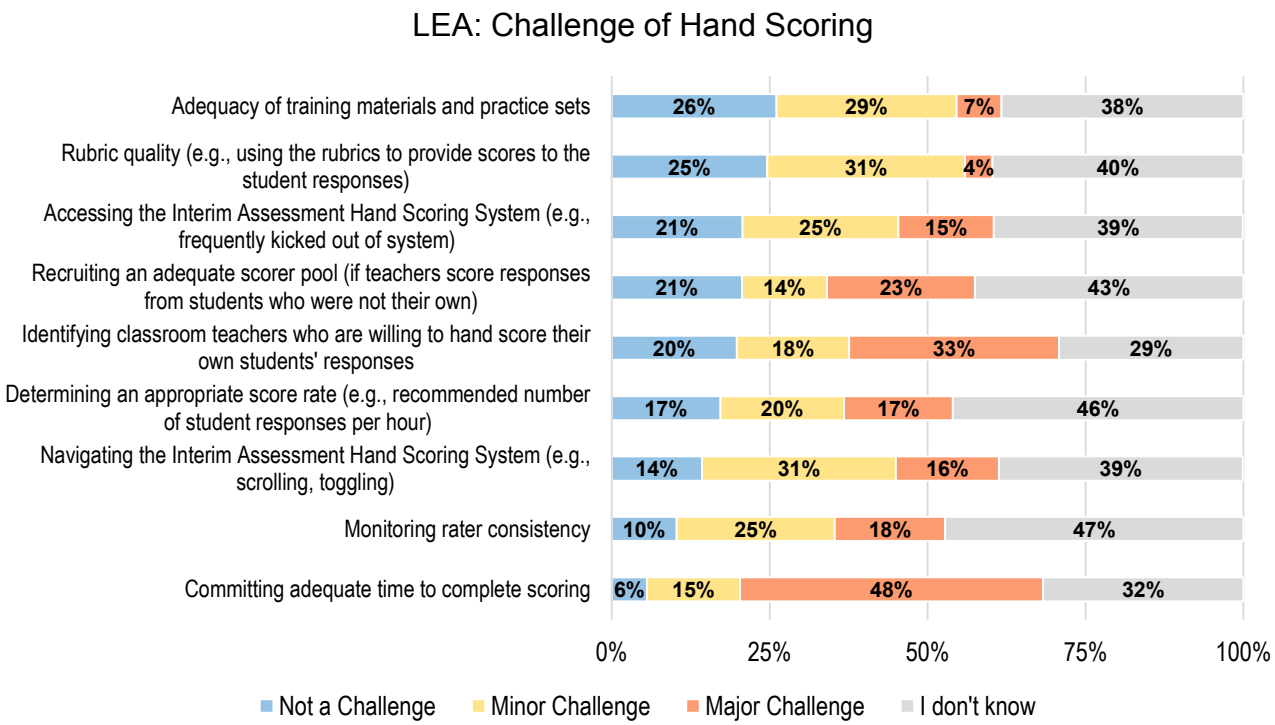


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.29. Challenge of hand scoring (LEA survey).

LEA coordinators indicated that adequacy of detail in reporting results (not reported by target CCSS or strand) to inform changes to instruction was the biggest reporting challenge (57% indicated it was a major challenge). This relates to LEA coordinators in our sample also perceiving the Interim Assessment Reporting System as being difficult to use because of the lack of detail offered. These findings reflect user perceptions absent an understanding of the psychometric limitations of reporting results based on very few test items.

Other major challenges included lack of integration with other student record systems (37%) and difficulty of aggregating/grouping student scores (32%) (see Figure 3.30).

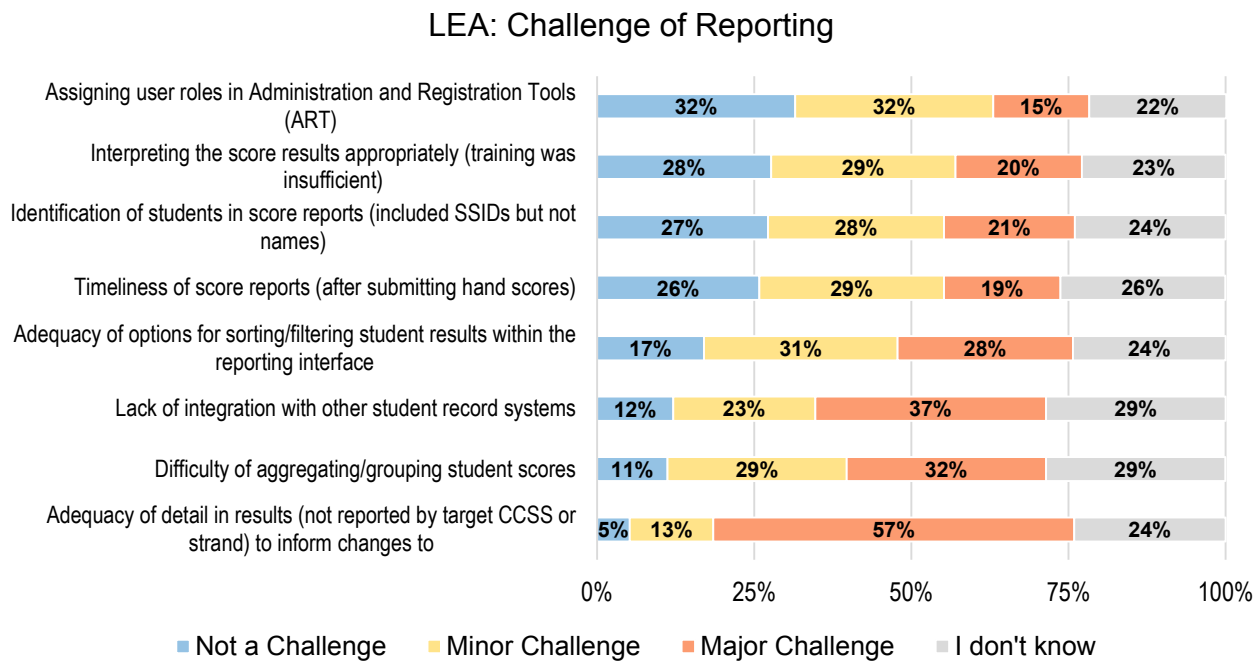


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.30. Challenge of reporting (LEA survey).

SITE Survey

The least challenging administration activity reported by school site coordinators was setting up the interim assessment for administration (62%). The most challenging activities were student familiarity with technology and scheduling time to access computers (22% indicated these as major challenges for both activities) (see Figure 3.31).

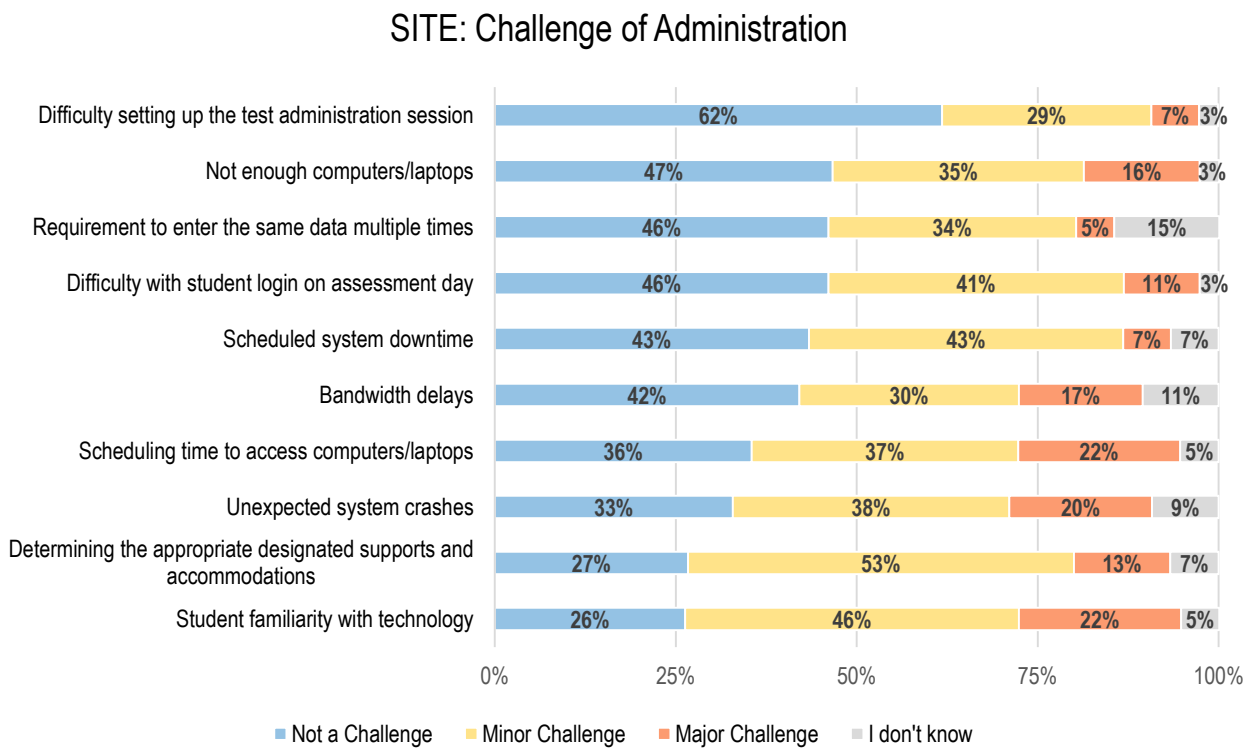


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.31. Challenge of administration (SITE survey).

Almost half of school site coordinators indicated the test content being too difficult was a major challenge (46%). Additionally, over one-third of the school site coordinators indicated students being confused by unfamiliar item types was a major challenge (37%) (Figure 3.32).

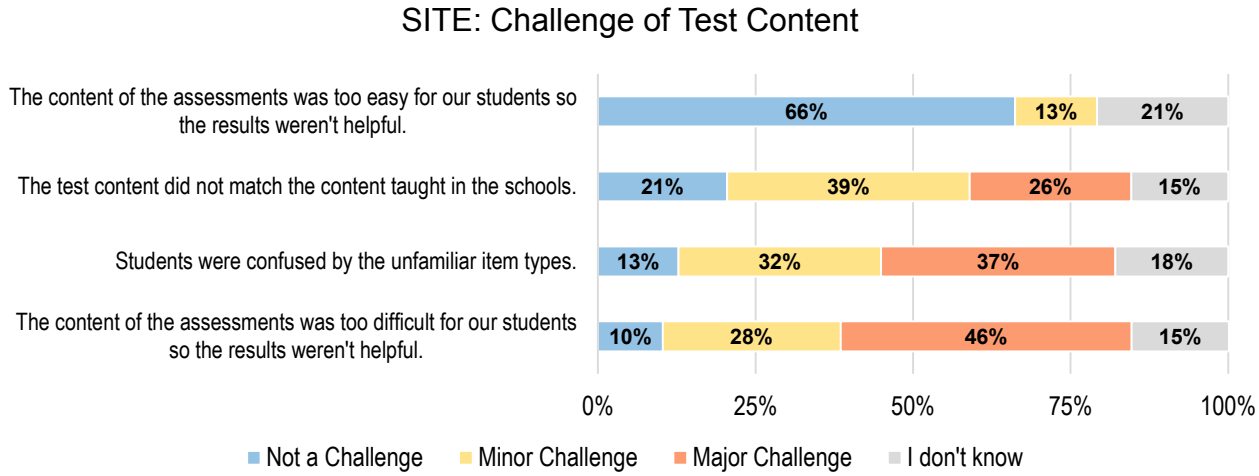


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.32. Challenge of test content (SITE survey).

Similar to LEA coordinators, school site coordinators reported committing adequate time to complete scoring was the greatest challenge with hand scoring activities (35.1% indicated it was a major challenge). Also, identifying classroom teachers willing to hand score their own students' responses was noted as a major challenge (32.4%). Many school site coordinators also reported not knowing whether some of the hand scoring activities were challenging or not (as indicated by "I don't know" ratings), suggesting they either did not participate in these activities and/or did not conduct any hand scoring (see Figure 3.33).

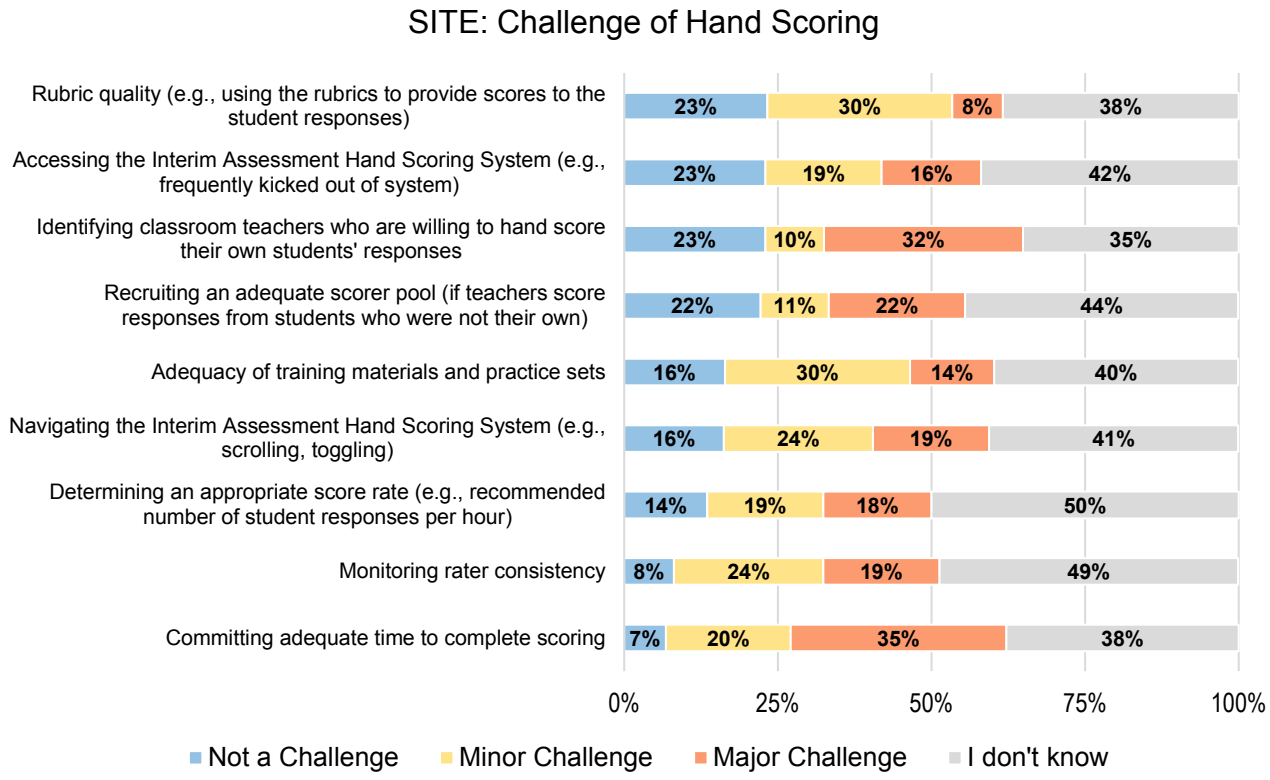


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.33. Challenge of hand scoring (SITE survey).

School site coordinators indicated adequacy of detail in reporting results (not reported by target CCSS or strand) to inform changes to instruction was the biggest reporting challenge (35.7% indicated it was a major challenge). Although school site coordinators specified other reporting activities were largely not challenging or presented only minor challenges, many did not know whether challenges existed or not, suggesting they had limited interaction with reports and reporting tools (see Figure 3.34).

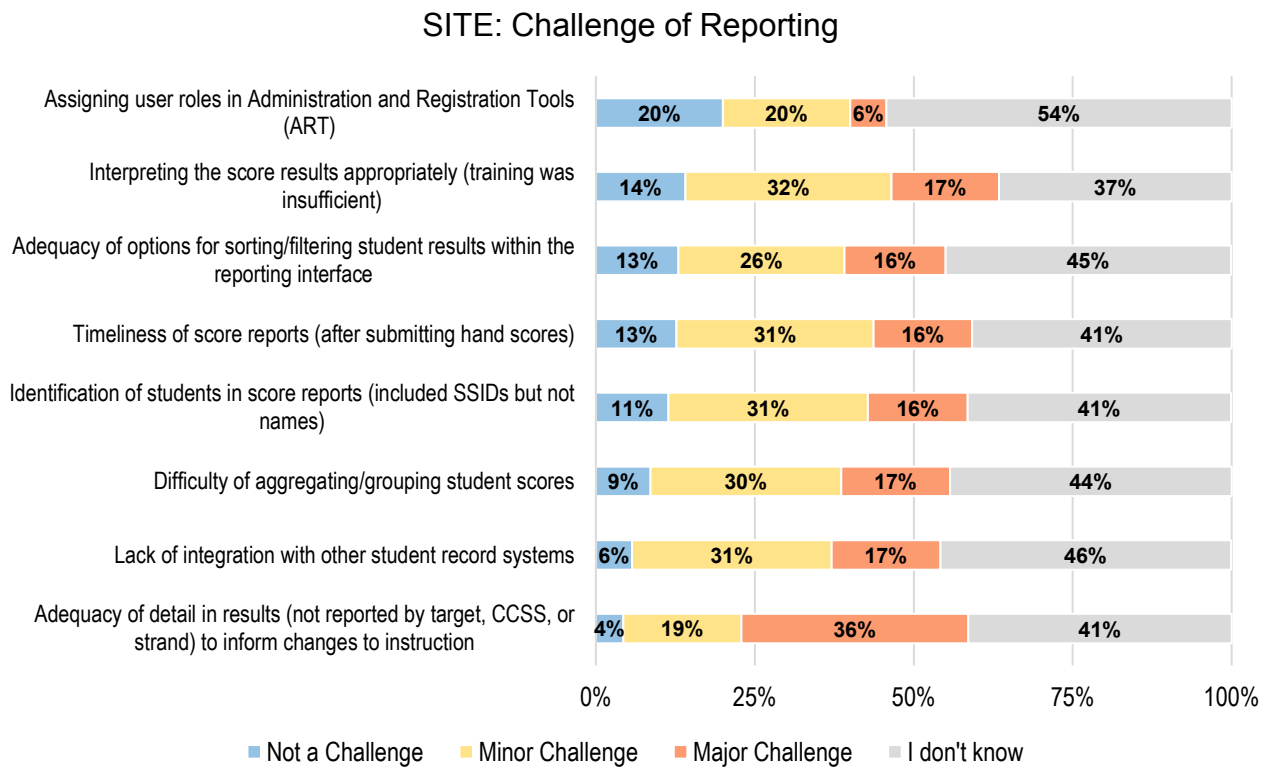


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.34. Challenge of reporting (SITE survey).

TA Survey

The least challenging administration activity reported by test administrators was setting up the interim assessment for administration (57.3% indicated this not a challenge). The most challenging activities were unexpected system crashes[29] (22.8% indicated this was a major challenge), scheduling time to access computers (22.6%), and student familiarity with technology (20.4%) (see Figure 3.35).

[29] The survey did not include probing questions to understand more deeply what the root cause of the system crashes were. That is, the system could have crashed because of issues with the Smarter Balanced test delivery system, local internet and networking issues, or some other related causes.

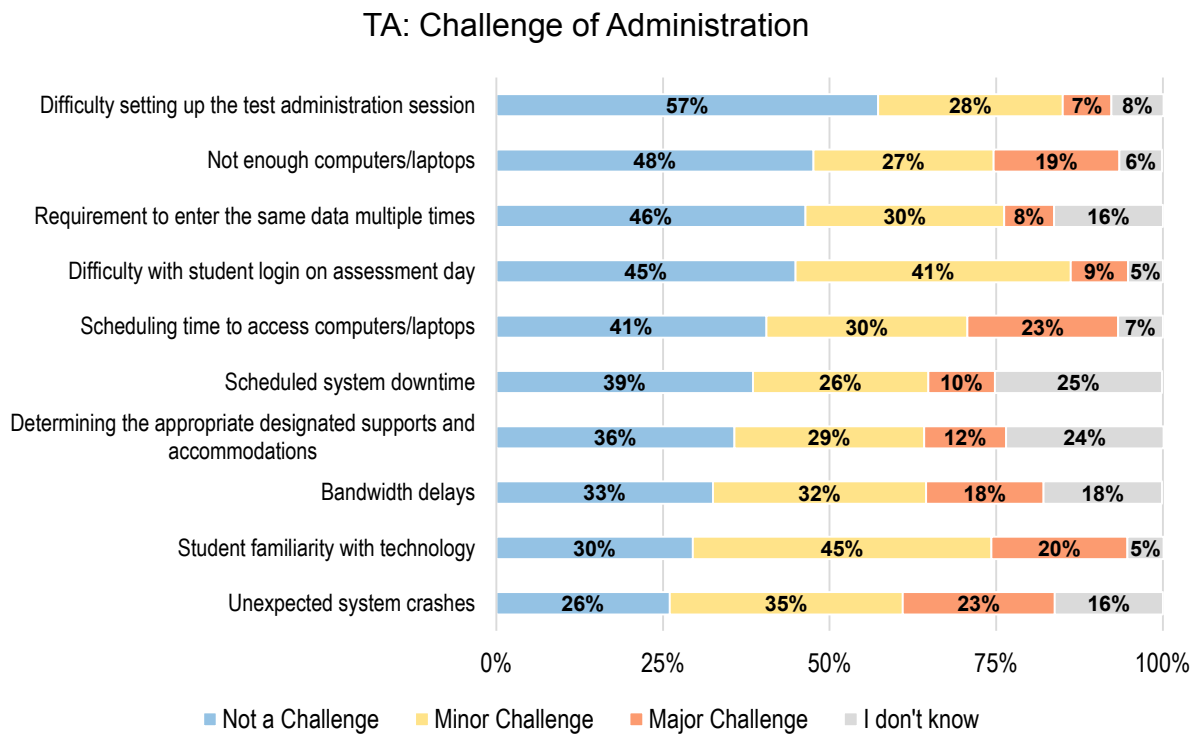


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.35. Challenge of administration (TA survey).

Almost half of test administrators indicated the test content being too difficult was a major challenge (47.2%). Additionally, over one-third of the test administrators indicated students being confused by unfamiliar item types also was a major challenge (35.7%) (see Figure 3.36).

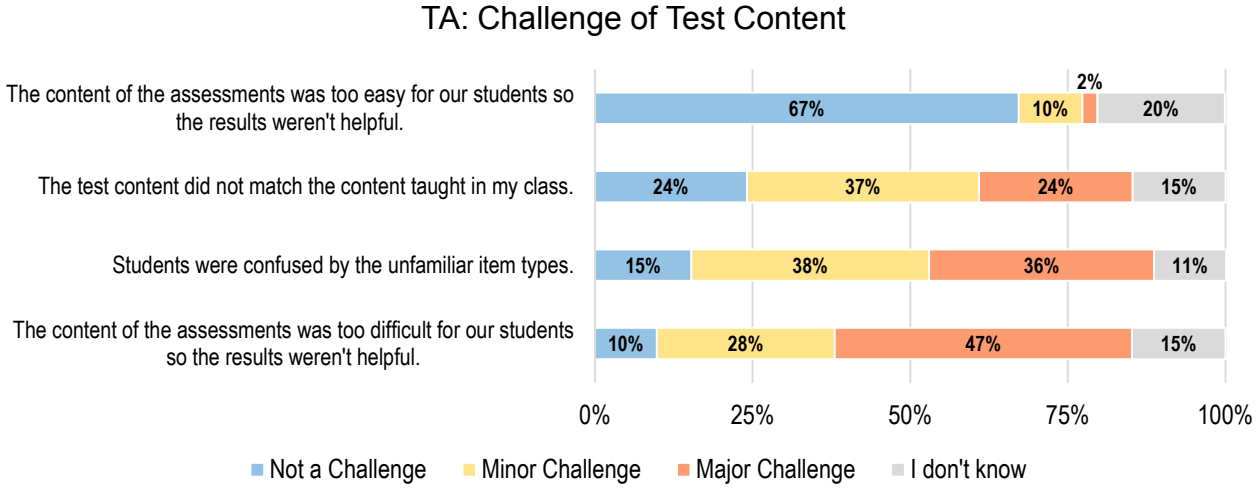


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.36. Challenge of test content (TA survey).

A very large number of test administrators in our sample reported they did not know whether various hand scoring activities were challenging, suggesting they might not have had any experience with hand scoring. Of the few test administrators who seemed to have some experience with hand scoring, they largely indicated the various activities were either not a challenge or only a minor challenge (see Figure 3.37).

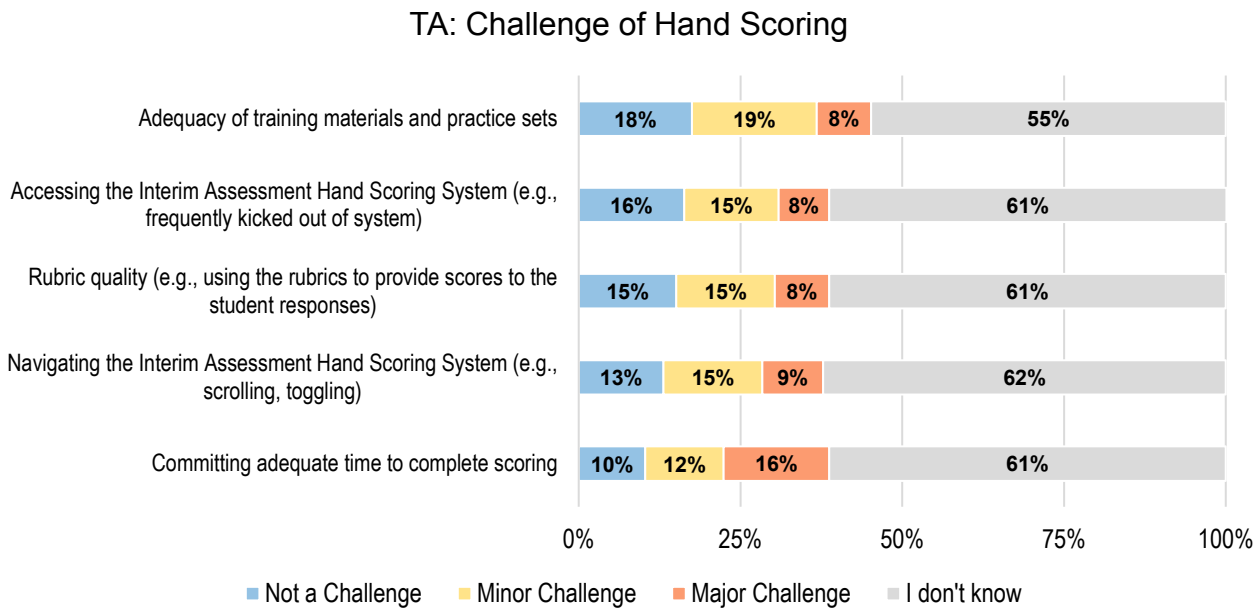


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.37. Challenge of hand scoring (TA survey).

For the most part, test administrators indicated they did not know whether challenges with reporting activities existed, suggesting they did not have much experience with these activities (see Figure 3.38).

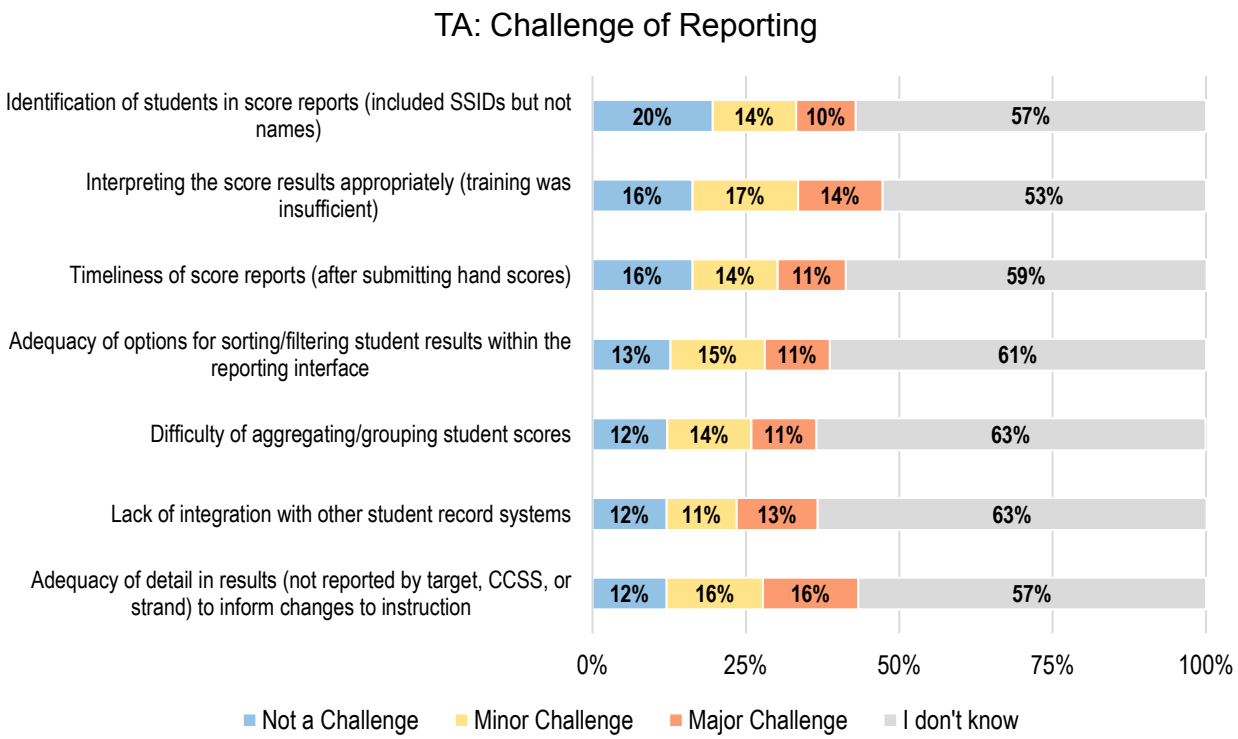


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.38. Challenge of reporting (TA survey).

Improvements

Table 3.26 shows the survey question numbers related to interim assessment improvement with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.26. Interim Assessment Improvement Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q20	78
SITE	Q24	79
TA	NA	NA

LEA and school site coordinators identified areas that could be improved for 2016–17. The two improvements most frequently cited as needed by LEA coordinators were to provide more detailed score reports (73.1%)[30] and to provide assessments that do not require hand scoring

[30] See earlier comments about the psychometric challenge of providing detailed score reports when reduced test length is also desired.

(66.3%)[31]. School site coordinators noted the same improvement needs (65.2% and 60.9%, respectively) (see Figure 3.39).

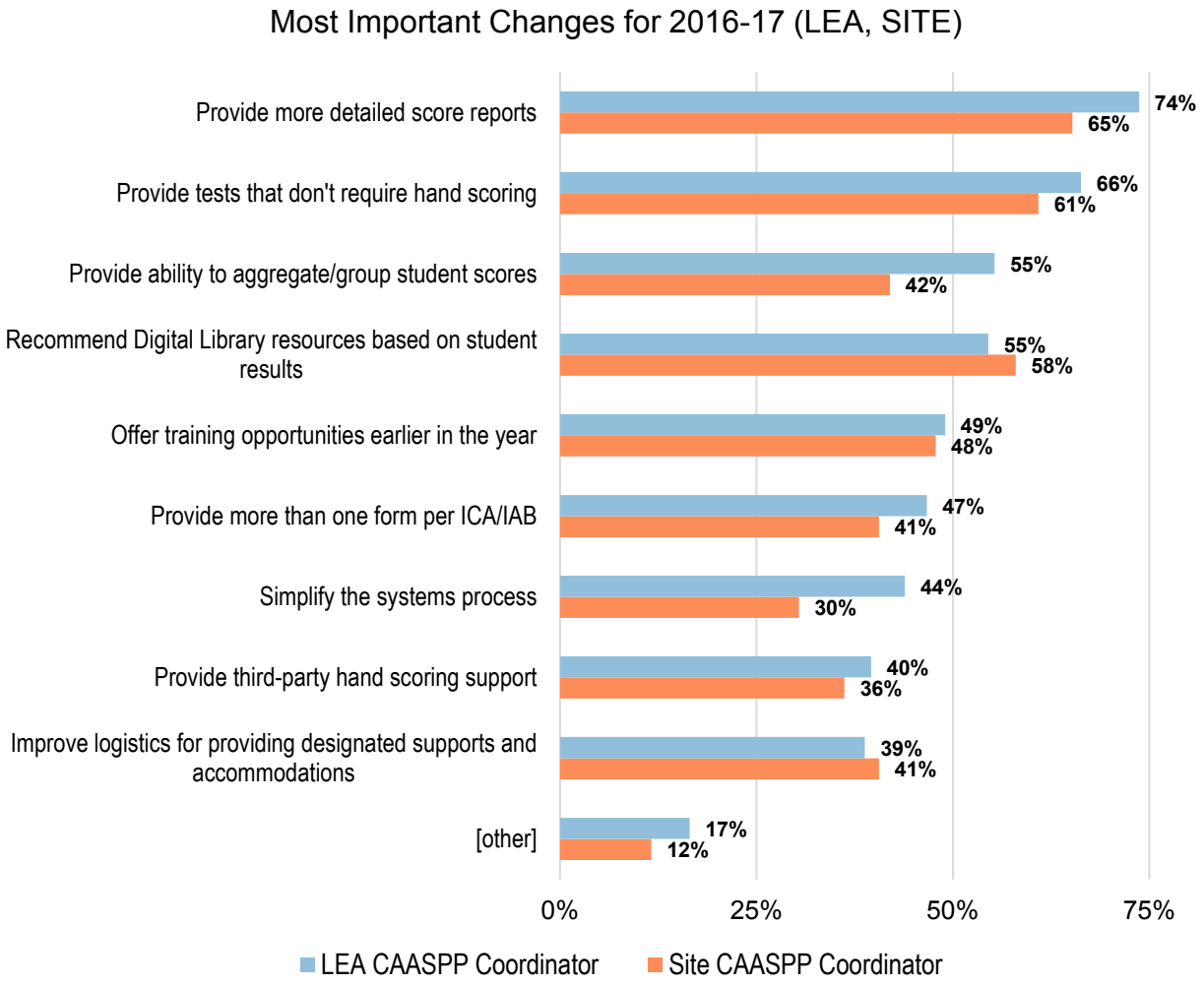


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 3.39. Most important changes for 2016–17 (LEA and SITE surveys).

2016–17 Plans

More participants in all three stakeholder groups indicated they would use the Smarter Balanced Interim Assessments in 2016–17 than said they would not permit or administer them. LEA coordinators generally indicated their LEA would either require the Smarter Balanced Interim Assessments (45.6%) in 2016–17 or highly encourage their use (13.1%). Because we heard overwhelming reports that hand scoring was a major factor in how and whether interim assessments were administered, we collected information regarding LEA interest in using external hand scoring contracting services in 2016–17, should CDE make it available. Only 8.7

[31] The survey question did not differentiate providing *more* tests that do not require hand scoring (IABs that do not require hand scoring are available) from providing *only* tests that do not require hand scoring; therefore, it is difficult to determine what respondents were referencing in this change.

percent of LEA coordinators reported they definitely would be interested. The large majority either indicated “no” (45.5%) or “maybe” (45.8%).

Approximately 60 percent of school site coordinators stated the interim assessments would be required and 29 percent stated their use would be highly encouraged in 2016–17. The test administrators were mixed in their responses; 47 percent stated they would administer the interim assessments, 29 percent indicated they would not, and 24 percent were not sure.

Additional Comments

Of the 285 LEA coordinators who administered Smarter Balanced Interim Assessments), 84 (or 29.5%) provided additional comments related to improvements to the interim assessments. Many individual responses touched on more than one issue. All comments were classified as issues with systems, training, design, scoring, results, or general positive feedback. The issues most commonly mentioned in the “additional comments” were “results” and “systems” issues. The most common response from LEA coordinators who provided additional comments was a desire to see more detailed results (n=32). Less commonly mentioned issues surrounded (a) improvement to the Digital Library (n=19), and (b) the need to provide third-party hand scoring support or find a way to make hand scoring less time consuming (n=11). LEA coordinators also provided a few generally positive comments, including (a) finding the interim assessment was an invaluable tool, (b) not experiencing any technology issues, and (c) stating the coordinator’s binder was extremely helpful. Full comment summaries are detailed in Appendix C12.

Only 16 of the 90 (17.8%) site coordinator respondents provided additional comments; however, those comments are discussed here because they represent specific issues with systems, training, design, scoring, or results. Full comment summaries are detailed in Appendix C12.

A total of 227 test administrators out of 863 (26.3%) provided additional comments related to improving the interim assessments. Again, many individual responses touched on more than one issue. The major issues identified were classified as issues with systems, training, design, scoring, or results. In addition, a small number of general positive comments were provided. Of the issues/concerns reported, the most frequent were related to “design” and “results.” Chief among all issues was a request to improve the questions and directions so they are clear to students (n=42). The remaining issues included (a) ensuring content is appropriate for SWD and EL students (n=27), (b) providing more detailed results (n=25), and (c) paring down the amount of text in questions and instructions (n=22). Test administrators also identified systems issues, including fixing errors/bugs (n=15), improving systems/Web site navigation and interface (n=16), and improving test navigation and interface (n=15). Three individuals noted the interim assessment was good practice for the summative assessment. See Appendix C12 for the remaining results.

Non-Usage

Small percentages (less than one-fourth) of each stakeholder group in our sample reported not administering any Smarter Balanced Interim Assessments in 2015–16 (Table 3.27). These non-users were asked a small subset of questions to provide insight as to why they did not use the Smarter Balanced Interim Assessments.

Table 3.27. Frequencies of Non-Interim Assessment Users by Survey

Survey	Frequency Not Administered	Percent
LEA	82	22.3
SITE	17	15.9
TA	181	17.3

Table 3.28 shows the survey question numbers related to interim assessment usage with their corresponding appendix table numbers for individual item descriptive statistics.

Table 3.28. Non-Use Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
LEA	Q26 – Q34	88, 91, 94, 97, 100 – 102
SITE	Q30 – Q39	10, 89, 92, 95, 98, 103 – 105
TA	Q27 – Q36	12, 90, 93, 96, 99, 106 –109

Policies

The majority of both LEA and site coordinators reported their LEA's policy for administering the interim assessments as voluntary with no specific encouragement (LEA: n = 51 of 82; SITE: n = 11 of 17). In contrast, only 43 of 181 test administrators reported that giving the tests was voluntary with no specific encouragement. Also noteworthy, slightly less than one-quarter of all stakeholders reported that their LEA prohibited interim assessment administration in 2015–16 (LEA: n = 19; SITE: n = 4, TA: n = 30).

Third-Party Assessments

To get a better understanding of how third-party assessments impact the use of Smarter Balanced Interim Assessments, we also asked nonusers what third-party assessment system they use. Over half of each group reported using a third-party interim assessment system (LEA: n = 61; SITE: n = 9; TA: 91). Similar to those who administered interim assessments, no single system was reported as being the most commonly used.

Factors Impacting Smarter Balanced Interim Assessment Use

Across all three surveys, no single reason why the interim assessments were not used dominated the responses. When asked to select only one reason, many LEA coordinators stated their currently used third-party assessment platform provides more benefits than does the Smarter Balanced Interim Assessment system (n = 19) and that teachers need to focus on other activities (n = 14). Test administrators cited their need to focus on other activities (n = 24) or said their LEA or school did not permit it (n = 22) as the primary reason for not administering. No clear patterns emerged for the site coordinators.

When asked to select any factor that significantly impacted the decision to administer Smarter Balanced Interim Assessments, test administrators most often cited the same primary factors;

32 stated that their LEA or school did not permit it and 32 stated they need to focus on other activities (Table 3.29). Approximately one-half of test administrators stated "other" factors (n = 72). These respondents noted as other factors that they weren't scheduled to administer the assessment (n=24), taught a subject area or grade level that was not assessed (n=22), or used a third party assessment (n=4).

Table 3.29. TA Primary Factors for Not Administering Interim Assessments

TA_Q30. What primary factors contributed to your decision to not administer the Smarter Balanced Interim Assessments in 2015-16? Select all that significantly impacted your decision.		
Response	**Frequency**	**Percent**
My LEA or school did not permit it.	32	21.6%
I need to focus on other activities (e.g., instructional time).	32	21.6%
Coordination and logistics were too difficult.	18	12.2%
Lack of resources (e.g., computers testing locations)	12	8.1%
Our currently used third-party interim assessment platform provides more benefits than the Smarter Balanced Interim Assessment platform.	7	4.7%
My LEA/school did not support hand scoring (e.g., stipends, professional development units).	6	4.1%
I decided that the Smarter Balanced Interim Assessments results would not provide enough actionable feedback.	5	3.4%
I wanted to wait and see how the Smarter Balanced Interim Assessments were perceived by others.	3	2.0%
[other]	72	48.6%
Frequency Missing = 30		

In addition to stating that teachers need to focus on other activities and citing their preference for their currently used third-party assessment system, many LEA coordinators also cited being unable to provide hand scoring support (n = 30) as a primary factor for not administering the interim assessments.[32] Table 3.30 indicates that LEA coordinators cited various reasons for not administering the interim assessments in 2015–16. Of the 15 LEA coordinators who indicated "other" factors that contributed to not administering the interim assessments in 2015–16, five of them stated training concerns (e.g., need more training, training not offered early enough in the year), three stated concerns about over-testing students, and eight indicated other reasons (e.g., administering the interim assessments would compromise too much instructional time, pressure from the teacher unions to not use the interim assessments).

[32] This would only impact administering IABs that require hand scoring or the ICAs.

Table 3.30. LEA Primary Factors for Not Administering Interim Assessments

LEA_Q29. What primary factors contributed to why the Smarter Balanced Interim Assessments were not administered in your LEA in 2015-16? Select all that significantly impacted your decision.		
Response	**Frequency**	**Percent**
Could not provide support for hand scoring (e.g. stipends, professional development units).	30	38.0%
Teachers need to focus on other activities (e.g. instructional time).	26	32.9%
Coordination and logistics were too difficult.	22	27.8%
Our currently used third-party interim assessment platform provides more benefits than the Smarter Balanced Interim Assessment platform.	22	27.8%
Our LEA decided that the Smarter Balanced Interim Assessment results would not provide enough actionable feedback.	19	24.1%
We wanted to wait and see how the Smarter Balanced Interim Assessments were perceived by others.	19	24.1%
Lack of resources (e.g. computers testing locations)	16	20.3%
[other]	15	19.0%
Frequency Missing = 3		

Some school site coordinators also cited hand scoring as a contributing factor to not administer the interim assessments (n = 6) (Table 3.31).

Table 3.31. SITE Primary Factors for Not Administering Interim Assessments

SITE_Q33. What primary factors contributed to why the Smarter Balanced Interim Assessments were not administered in your school in 2015-16? Select all that significantly impacted your decision.		
Response	**Frequency**	**Percent**
Could not provide support for hand scoring (e.g., stipends, professional development units).	6	40.0%
Teachers need to focus on other activities (e.g., instructional time).	5	33.3%
Coordination and logistics were too difficult.	3	20.0%
Our LEA decided that the interim assessment results would not provide enough actionable feedback.	3	20.0%
Lack of resources (e.g., computers, testing locations)	2	13.3%
Our currently used third-party interim assessment platform provides more benefits than the Smarter Balanced Interim Assessment platform.	2	13.3%
Our LEA made the decision.	2	13.3%
Our LEA did not permit it.	1	6.7%
We wanted to wait and see how the Smarter Balanced Interim Assessments were perceived by others.	1	6.7%
[other]	2	13.3%
Frequency Missing = 2		

Plans for 2016–17

The majority of LEA and school site coordinators who did not use the interim assessments in 2015–16 stated they expect the Smarter Balanced Interim Assessments to be voluntary (highly encouraged or no specific encouragement) in 2016–17 (LEA: n = 55; SITE: n = 12); smaller percentages stated they expected the interim assessments to be required in 2016–17 (LEA: n = 16; SITE: n = 0). In contrast, a higher percentage of TAs stated they expect the Smarter Balanced Interim Assessments to be required in 2016–17 (n = 59).

When asked why the interim assessments will be allowed in 2016–17, many LEA coordinators selected the response "We have heard teachers and students found that practice with the system and exposure to item types was helpful for administering the summative assessments" (n = 31). Some also selected the response "We have heard that the interim assessments provided actionable feedback about student learning" (n = 16). See Table 3.32 for full results.

Of the 19 LEA coordinators who indicated "other" factors, 17 noted specific reasons we had not asked about. The primary reason mentioned by most LEA coordinators (n = 8) was the idea of teacher choice—that providing the interim assessments as a resource allowed teachers to decide whether or not to implement them. Three individuals specifically noted that they thought the students would benefit from the practice. There were some qualifiers given as to whether or not the interim assessments would be allowed in 2016–17. Hand scoring was a concern from three individuals and one mentioned other issues to be worked out. Finally, two individuals indicated that they were considering the interim assessments in lieu of district benchmark assessments and want to try them out.

Table 3.32. LEA Primary Reasons for Allowing Interim Assessment Use in 2016–17, of Those That Did Not Use Them in 2015–16

LEA_Q31. What are the primary reasons for allowing or requiring the Interim Assessments in 2016-17? Select all that significantly impacted this decision.		
	Frequency	Percent
We have heard that teachers and students found that practice with the system and exposure to item types was helpful for administering the summative assessments.	31	47.7%
We have heard that the interim assessments provided actionable feedback about student learning.	16	24.6%
We are able to better support training, administration and preparation activities due to additional staffing or funding.	9	13.8%
We have heard that teachers benefited from conducting the hand scoring activities.	8	12.3%
Our educators have requested they be allowed to use it.	7	10.8%
We have heard that administering the interim assessments went smoothly.	4	6.2%
Our third-party assessment system contract is expiring.	1	1.5%
[other][a]	19	29.2%

Some school site coordinators also reported hearing that teachers and students found that practice with the system and exposure to item types was helpful for administering the summative assessments (n = 4). See Table 3.33 for full results.

Table 3.33. SITE Primary Reasons for Allowing Interim Assessment Use Next Year, of Those That Did Not Use Them in 2015–16

SITE_Q35. What are the primary reasons for allowing or requiring the Interim Assessments in 2016-17? Select all that significantly impacted this decision.		
Response	**Frequency**	**Percent**
We have heard that teachers and students found that practice with the system and exposure to item types was helpful for administering the summative assessments.	4	40.0%
We have heard that teachers benefited from conducting the hand scoring activities.	3	30.0%
Our educators have requested they be allowed to use it.	3	30.0%
We have heard that the interim assessments provided actionable feedback about student learning.	2	20.0%
We have heard that administering the interim assessments went smoothly.	1	10.0%
Our third-party assessment system contract is expiring.	1	10.0%
We are able to better support training, administration, and preparation activities due to additional staffing or funding.	1	10.0%
[other]	3	30.0%
Frequency Missing = 4		

Additional Comments

Many of the LEA CAASPP coordinators from LEAs that did not administer the Smarter Balanced Interim Assessment (n = 18) provided additional comments. More than one comment related to difficulty hand scoring items (n = 4), the results of the assessments not being helpful because they are not detailed enough or at the item level, and the amount of time it takes to administer the assessment. Individual comments indicated that one LEA was piloting the assessments next year, there is a need for training early in the school year, there should be items specific to grades nine and ten, the training materials provided were insufficient, there should be a contingency for students who tested after the rest of the group, there's no time to reteach content even if areas of weakness are found, and that it would be helpful to be able to save the designated accommodations for students to avoid having to enter the same information multiple times.

Only three school site coordinators who did not use the interim assessments provided additional comments. One individual expressed concern about the length of the *Summative Assessments*. Another individual mentioned the time commitment to hand scoring and that there are several other ways to gather similar information without hand scoring. The final respondent stated the belief that the test was inappropriate for students with moderate to severe cognitive disabilities.

Thirty-one test administrators provided additional concerns, comments, or ways to improve the Smarter Balanced Interim Assessments. The major issues identified were with systems, design, scoring, and results. Of these, design issues were most frequently discussed (n=7). The most frequent individual issues included the need to simplify the testing process (n=4), and a need to better incorporate accommodations/resources in the design to better serve SWDs (n = 3). There also were a few positive comments, including feeling the practice assessments were easily

accessible, appreciating the ability to practice navigation, and plans to use the interim assessments next year.

Additional Analyses of Survey Response Data

In addition to conducting the descriptive analyses of each survey, we further analyzed survey data by comparing responses to several questions by different subgroups of respondents. The following variables were used to generate cross tabulations:

- LEA student population (large or small)
- LEA type (charter or district)
- LEA EL population (high or low)
- Subject and grade level the test administrator primarily taught

For some of these cross tabulations, the number of responses in one or more subgroups was too low to be reliable or informative. For a number of the survey questions, the cross tabulations did not indicate notable differences in responses by subgroup. However, within our analytic samples, several cross tabulations provided additional insight into relationships between local experiences with the interim assessments and (a) LEA size and (b) whether or not an LEA is a charter, despite the limitation of being able to generalize to all subgroups within the state. Specific findings regarding the cross tabulations are provided in chapter 6.

Analysis of Statewide Testing Data

In addition to the results from the focus groups and statewide survey, it is also instructive to examine testing data. While we do not have access to score data for the interim assessments, ETS provided a file indicating the use of interim assessments for the state population. The interim assessments were administered to a large number of California students. In the 2015-16 school year, 4,321,412 interim assessments were administered through March 2016 (some students took multiple interim assessments). The IABs were administered much more frequently than the ICAs. A total of 3,439,424 IABs were administered, compared to 881,988 ICAs.

When we examine the use of the interim assessments at the school level, 6,178 schools administered at least one ICA or IAB assessment, compared to 3,034 that did not administer either. Table 3.34 presents the number of schools that used and did not use the interim assessments (only schools with a scale score for 2016 greater than 1 on the statewide data system[33] are included in Table 3.34). We see that more schools administered IABs than ICAs. The great majority of schools that used ICAs also used IABs. Table 3.34 also shows that schools that used the interim assessments had lower aggregated Smarter Balanced Summative Assessment English language arts/literacy (ELA) and mathematics scale scores than those who did not. The mean differences ranged from about 10 to 15 points.

[33] Downloaded from http://caaspp.cde.ca.gov/sb2016/ResearchFileList

Table 3.34. Interim Assessment Usage and Mean Scale Score (2015–16)

Interim Assessment Type	Interim Assessment Status		Summative Assessment Scores	
	Usage	N Schools	Mean	Difference (Not used – Used)
IAB	Not Used	3,363	2496.76	
	Used	5,840	2486.21	10.55
ICA	Not Used	5,464	2495.98	
	Used	3,739	2481.42	14.56
IAB and ICA	Not Used	5,800	2495.15	
	Used	3,403	2481.40	13.75

Note: Scale scores for the Smarter Balanced Summative Assessment for English language arts/literacy (ELA) and mathematics are aggregated. Only non-zero scores are included.

Table 3.35 presents the school-level gains from 2015 to 2016. While those using interim assessments tended to have lower mean scale scores in 2016, they had higher change (or gain) scores. Schools that used the IAB, ICA, or a combination of the IAB and ICA interim assessments had greater increases in their aggregated Smarter Balanced Summative Assessment English language arts/literacy (ELA) and mathematics scale scores between 2015 and 2016. Regardless of the interim assessments(s) being used, the difference between the mean scale score changes for schools that used and did not use the assessments is quite small (IAB: 0.29 points; ICA: 0.67 points; IAB and ICA: 0.52 points). It should be noted, however, that each of these scale score changes are statistically significant (p <.001 for each assessment) and in the direction we would hope, showing greater improvement for the schools that are using the ICAs or IABs.

Table 3.35. Interim Assessment Usage and Mean Scale Score Change (2015–2016)

Interim Assessment Type	Usage	N Schools	Mean	Std. Deviation	Significance (2-tailed)
IAB	Not Used	3,213	7.88	23.48	*P* <.001
	Used	5,774	8.17	19.49	*P* <.001
ICA	Not Used	5,284	7.79	21.10	*P* <.001
	Used	3,703	8.46	20.86	*P* <.001
IAB and ICA	Not Used	5,614	7.87	20.85	*P* <.001
	Used	3,373	8.39	21.26	*P* <.001

Note: Scale scores for the Smarter Balanced Summative Assessment for English language arts/literacy (ELA) and mathematics are aggregated.

CHAPTER 4: ACCESS TO DESIGNATED SUPPORTS AND ACCOMMODATIONS STUDY

Emily Dickinson and Richard Deatz

Background

One of the underlying assumptions of the CAASPP system is that the test results provide accurate and unbiased assessments of what all students have learned and what the gaps in learning are as well. To obtain such results, it is important that all students have full and equal access to the test in order to be able to demonstrate learning. Aspects of the test that might impede student access to the test content or that impede students' ability to demonstrate what they know and can do threaten the accuracy and validity of test results.

As part of HumRRO's Independent Evaluation Plan, which was approved by the California Department of Education (CDE) and the State Board of Education (SBE), we were charged with conducting a study during the 2015–16 school year to explore student access to CAASPP assessments. The specific purposes of this study are to (a) examine the availability and use of testing supports and accommodations for students with disabilities (SWDs) and English learners (ELs) on the Smarter Balanced ELA and mathematics assessments and (b) determine whether the tools used for these assessments are consistent with those used routinely by classroom teachers and their students.

This chapter presents the design, data sources, and results of the Access to Designated Supports and Accommodations Study. The study relies on two primary sources of data: (a) information generated by volunteers from HumRRO's Local Educational Agency Research Network (LEARN[34]), and (b) extant data on the provision of testing accommodations during the 2016 Smarter Balanced Summative Assessments. The results from this small study provide insights into how instructional practices used in the classroom with SWDs and ELs relate to the currently available universal tools, designated supports, and accommodations offered during online Smarter Balanced Summative Assessments.

Literature Review

Standard 3.9 of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) states: "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" (p. 67). Multiple types of disabilities that might impact obtaining a valid test score under typical administration have been defined. Across the testing literature, more than 70 accommodations have been identified to remove construct-irrelevant barriers to performance (Olson, 2010). According to Sireci (2004), accommodations are meant "to level the playing field so that the format of the test or the test administration conditions do not unduly prevent such students from demonstrating their 'true' knowledge, skills, and abilities" (p.1). Common accommodations include extended time, oral presentation and response, braille, sign language interpreters, format changes (e.g., large print), relocation to a quiet room, and computerized aids (Weston, 2003).

[34] See Chapter 2 for a full description of the LEARN.

Assessment accommodations provided to SWDs should mimic what is provided in the classroom, but not if they modify the construct. Elliott, Thurlow, Ysseldyke, and Erickson (1997) note that “accommodations should be provided for the assessment when they are routinely provided during classroom instruction. In other words, when classroom accommodations are made so that learning is not impeded by a student's disability, such accommodations generally should be provided during assessment” (p.4). The National Center for Learning Disabilities (Crawford, 2007) examined the test accommodation policies of six states (California, Colorado, Massachusetts, Michigan, Texas, and Wyoming) and found that all stressed the importance of using accommodations in the classroom prior to being provided as a test accommodation, yet only two of the states (Colorado and Wyoming) required that the accommodation was available to the student at least 90 days prior to test administration with evidence the accommodation was effective.

Various studies have been conducted to examine the validity of using accommodations in standardized testing. Sireci, Li, and Scarpati (2003) reviewed 150 accommodations studies to examine the interaction between accommodation condition and type of student (i.e., whether SWDs would see improved standardized test scores with accommodations while general education students would not). Overall, their review found inconsistent results; however, they consistently found that the extended time accommodation increased test performance of SWDs, whereas general education students who were provided extended time accommodation did not improve comparably. When conducting a similar review on the validity of accommodations, Thurlow (2007) found the use of oral accommodations consistent with Sireci et al (2003); she found a positive impact occurred in approximately half the studies of mathematics assessments but no impact in other content areas, while the positive impact of extended time remained consistent. Weston (2003) found the use of oral accommodations had a significant impact on test performance for SWDs with the lowest reading abilities, with more mixed results for those SWDs with higher reading abilities. This suggests the possibility that the inconsistent findings noted by Thurlow (2007) and Sireci et al (2003) might have been a result of not distinguishing between low and high ability readers.

The Connecticut Enhanced Assessment Grant (CTEAG) was implemented to examine the validity of inferences from accommodated tests (Olson, 2010). Four states participated in at least one study, with each implementing research designs that assigned students with general and special education designations to accommodation and non-accommodation testing conditions. Each state selected the accommodation(s) to study, with read-aloud the most frequently selected. The studies generally found that test accommodations did not impact item statistics and test characteristics; in most cases, the accommodations increased the scores of SWDs significantly more than those of students in the general population.

Providing accommodations to SWDs is not only desirable for validity purposes, but also specified in the US peer review guidelines. Based on these guidelines, states need to show that tests are accessible to all students by providing a variety of accommodations and that the use of accommodations yields meaningful scores. In addition, these guidelines require evidence that the accommodations provided for the assessment are consistent with accommodations that students receive during instruction and testing in the classroom (Thurlow, 2007).

The present study adds to this literature base by taking a closer look at the extent to which the summative testing supports or accommodations offered to SWDs and ELs in California correspond to the instructional practices that they are exposed to on a regular basis.

CAASPP Universal Tools, Designated Supports, and Accommodations

The CAASPP allows universal tools, designated supports, and accommodations (embedded and non-embedded). ***Universal tools*** are available for all students based on student preference and selection, and include, among other resources, breaks, digital notepad, English glossary, highlighter, strikethrough, zoom, and mark-for-review. ***Designated supports*** are available to all students when determined for use by an educator or group of educators (with parents/guardian and student input, as appropriate) or specified in the student's individualized education program (IEP) or Section 504 plan. These include but are not limited to color contrast, masking, text-to-speech, ability to turn off any universal tool, separate setting, special lighting or acoustics, and administration of the test to the pupil at the most beneficial time of day. ***Accommodations*** are available if specified in the student's IEP or Section 504 plan and include braille; text-to-speech for ELA reading passages; alternate response options; American Sign Language for Writing, Listening, and Mathematics; print on demand; read aloud for ELA reading passages; scribe for Writing; and speech-to-text.[35]

Study Design

Research Questions

To meet the purposes of the study, we addressed the following research questions:

1. Is the general assessment accessible to moderately disabled students[36] and ELs through the provision of accommodations and supports?
2. To what extent do the supports and accommodations provided and used in the interim and summative assessments match those used in classroom instruction for individual students?
3. Are there types of supports or accommodations used by students when learning in the classroom that are not used on assessments?
4. How often do students attempt to use test supports and accommodations that they do not use during classroom instruction?

Data Sources

Generated Data

We generated qualitative and quantitative data to address the study's research questions. The qualitative data were generated from phone-based focus groups, in-person interviews with local staff during school site visits, and observations of instruction and assessment. Quantitative data related to the use of instructional accommodations were generated from teacher reflection checklists and from items on the survey conducted as part of the Interim Assessment Administration Study (see Chapter 3). The processes for collecting these data are described in the Methods section of this chapter, below.

[35] Source: http://www.cde.ca.gov/ta/tg/ai/caasppmatrix1.asp (Retrieved 10/5/16).
[36] Moderately disabled students as defined by CDE. This includes students who are not eligible for the California Alternate Assessments (CAAs).

Extant Data

Data on the provision of testing accommodations during the Smarter Balanced Summative Assessments are maintained by the Educational Testing Service (ETS), the assessment contractor for the CDE. HumRRO worked in cooperation with CDE and ETS staff to obtain these data. These data were merged with the generated data on instructional practice use described above.

Methods and Analysis

This section first describes the process to recruit study participants and the characteristics of the final research sample. This information is followed by descriptions of the data sources, including development and implementation of the data collection protocols and analysis of the data.

Recruitment

As described in Chapter 2, we obtained commitments from 24 LEAs to join our LEARN, which we intended to function as a collaborative network on various aspects of the evaluation studies including participation in data collection activities. During the inaugural meeting with LEARN representatives in December 2015, we introduced the Access to Designated Supports and Accommodations Study, outlined the study plan, and provided a flier describing the study and the anticipated time commitments (see Appendix D1). This was the start of the recruitment effort, which occurred across multiple stages. First, we recruited LEAs to participate in the study. Next, we engaged LEA representatives to assist in recruiting other district- and school-level staff. Finally, district- and school-level staff assisted in recruiting additional school-level participants.

At the end of the study commitment period, only five LEAs had indicated interest and willingness to participate. The Study Director proceeded to the next stage of recruitment and contacted each of these volunteer LEA representatives directly, providing each with more information and a request to develop a recruitment plan for local staff within the LEA. During this time, one of the five initial LEAs dropped out of the study. HumRRO and CDE staff discussed participation levels and determined that the originally targeted sample size could be reduced (from 12 LEAs to 8).

To augment the small sample, we embarked on additional unplanned efforts to recruit more LEAs. CDE staff contacted select LEAs that were part of the LEARN to encourage their participation. CDE staff also identified seven LEAs outside of the LEARN and made introductions for HumRRO staff to recruit them. We focused on contacting “non-LEARN” LEAs with substantial EL populations to ensure the study included perspectives on issues related to designated supports and accommodations among ELs.

To maintain the planned study timeline, we continued data collection efforts among the four initial LEAs while also pursuing additional recruitment efforts. These efforts resulted in recruiting two additional LEAs. However, because we continued planned activities with the four initial LEAs, these two late-joining LEAs did not participate in the focus group component of the study. Table 4.1 presents descriptive information on the final sample of LEAs participating in the study.

The Study Director contacted the LEARN representatives of LEAs confirmed to participate and reminded them of the following key activities relevant to internal recruitment efforts:

- Identify approximately three schools within the LEA. Ideally, each school would represent a different grade span (e.g., elementary, middle, high).

- Within each school, identify one to four staff (e.g., teachers, content specialists, assessment coordinators) who routinely provide ELA and/or mathematics instruction and/or administer assessments to SWDs and ELs.
- Identify additional staff at the LEA level knowledgeable about using accommodations and supports for instruction and testing to participate in the focus groups/interviews.

We presented two options to the LEAs to support their recruitment of local staff. The first option required LEARN representatives to provide us with a list of all schools (including school-level demographic information) and a school point of contact (POC). HumRRO would identify a sample of schools from that list, notify the LEARN representative, and if acceptable, contact the school POCs to confirm and further recruit individual staff. The second option required that LEARN representatives identify the school sample and provide us with contact information for school POCs and/or interested staff. All participating LEAs chose the second option. HumRRO worked with school-level POCs to establish dates and times that would be most convenient for staff while also allowing HumRRO to observe both instruction and assessment. Table 4.2 presents descriptive information about the final sample of schools participating in the study.

Table 4.1. Characteristics of LEAs Participating in the Access to Designated Supports and Accommodations Study (N=6)

	LEA Characteristics			
Size	Small	Medium	Large	Very Large
n	1	3	1	1
% EL	Low	High		
n	1	5		
% SWD	Low	High		
n	4	2		
Academic Index	Average or lower	Above average		
n	6[37]	0		
Region	South	North		
n	2	4		

Notes: Size is based on number of students with 2015 CAASPP records (small= less than 1,000; medium= 1,000–8,999; large= 9,000–34,999; very large= 35,000 or more. Low percentage of ELs= 0%–15%. High percentage of ELs= more than 15%. Low percentage of SWDs= 0%–12%. High percentage of SWDs= more than 12%. Academic index was computed by computing grade-level means for each test to z-scores and then computing a weighted average of the z-scores across grades and subjects (weighted by the number of students with data for each grade and subject). South region= zip codes below 93000. North region = zip codes 93000 and higher.

[37] An attempt was made to find additional districts with a significant number of ELs and SWDs. As a consequence, nearly all of the LEAs considered had low (below state average) academic performance as indicated by 2015 Smarter Balanced scores.

Table 4.2. Characteristics of Participating Schools (N=15)

	Mean (SD)	Median	Minimum	Maximum
Enrollment Size	1,143 (790)	858	75	2,532
% Economically Disadvantaged	80.6% (16.0%)	85.6%	37.8%	94.7%
% Students with Disabilities	12.3% (14.7%)	7.8%	3.9%	61.5%
% English Learners	28.0% (21.6%)	20.0%	1.9%	80.9%
% Proficient ELA	29.5% (15.1%)	25.0%	11.0%	65.0%
% Proficient Math	17.9% (14.3%)	16.0%	4.0%	56.0%

Notes. Table represents data from 2014–15 School Report Cards. Economically disadvantaged is defined by California as students who are included in the national school lunch program or whose parent's educational attainment was less than a high school diploma. For all schools, the primary language of ELs was Spanish.

The final sample was smaller than planned across all stakeholder groups included in the study. The six participating LEAs represent 50 percent of the originally planned sample and 67 percent of the sample included in the modified sampling plan determined in cooperation with the CDE. As noted, not all LEAs participated in the focus group component of the study. The numbers of schools and staff relative to the number of LEAs were closer to targets, though they still fell short of the three schools and 6–10 staff per LEA. Because of the smaller sample sizes, we were able to include more qualitative analyses, leading to a deeper understanding of supports and accommodations used in each school, though this greatly limits the generalizability of the results. Table 4.3 shows a comparison of the sampling targets and the final research sample.

Table 4.3. Planned Sample vs Research Sample

	Number of LEAs	Number of Schools	Number of Staff
Original Plan	12	36	72–120
Modified Plan	8	24	48–80
Final Sample	6	15	34

Focus Groups and Interviews

Protocol Development and Implementation

We developed focus group protocols based on a review of materials available on the CAASPP Web site, along with materials gathered by HumRRO staff when participating in CAASPP training workshops. HumRRO staff drafted initial questions, reviewed and finalized questions, and created a final format that allowed a balance between consistency across users and flexibility to address unique topics as they arose.

The protocols included guidance for HumRRO facilitators. This guidance included opening and closing scripts, reminders of the overarching research questions, and background information that provided a rationale for particular questions. Protocols also recommended probing questions and indicated to which stakeholder group each question should be posed (e.g., LEA staff, instructional staff, both). The final focus group protocol is presented in Appendix D2.

Focus groups were conducted via telephone. We contacted school and LEA staff identified by the LEARN representative and asked them to indicate their availability via Doodle Poll. When multiple staff could not be scheduled for one meeting, we conducted individual interviews. Each focus group/interview was facilitated by one HumRRO staff person while a second staff member recorded notes. Both HumRRO staff reviewed the notes to create a final summary for analysis.

Analysis

Focus group/interview data were analyzed at multiple stages. First, we analyzed data to determine preliminary results to inform the remaining data collection components. We shared these initial results with LEARN representatives during the spring 2016 LEARN meeting, described in Chapter 2, to obtain their feedback on the findings and to elicit additional information. Notes from this meeting were appended to the initial focus group/interview data for inclusion in subsequent analyses.

Initial analysis of focus group/interview data consisted of hand coding notes to identify themes. Two HumRRO staff members identified meaningful categories that emerged from multiple reviews of the notes. Data from the focus groups/interviews were further analyzed in conjunction with school site visit data. This process is described in detail in a subsequent section.

School Site Visits

HumRRO staff conducted school site visits including multiple components during April and May, 2016. During each school visit, we attempted to conduct the following:

- Interview with the school's site CAASPP coordinator.
- Interviews with classroom teachers who provide instruction and/or administer assessments to SWDs and/or ELs.
- Observations of one or more classroom periods during which ELA or mathematics instruction was conducted.
- Observations of assessment administration (formative, interim, or summative)

Schools varied in terms of the site visit components that HumRRO staff was able to conduct. In some schools, the visits occurred during the summative assessment window and some teachers were in the process of conducting the Smarter Balanced Classroom Activity rather than a regular lesson. In one school visited, staff stated the school did not have an on-site CAASPP coordinator, while in other schools CAASPP coordinators were unable to be interviewed due to their testing-related responsibilities. Table 4.4 presents the total number of site visit components completed across the 15 schools visited.

Table 4.4. Number of Site Visit Components Completed (N Schools= 15)

Site Visit Component	Number Completed
CAASPP Coordinator Interviews	10
Teacher/Test Administrator Interviews	24
Classroom Observations	17
Assessment Observations	12

Protocol Development and Implementation

Several protocols were developed to capture data from each component of the school site visits. Protocol development was guided by preliminary findings from the focus groups, along with information and materials available on the CAASPP Web site. Final versions of the protocols are presented in Appendix D2 and D3.

Site visitors brought several resource documents with them to support the interviews and observations, including:

- *Smarter Balanced Assessment Consortium: Usability, Accessibility, and Accommodations Guidelines*
- *Matrix One: Universal Tools, Designated Supports, and Accommodations for the California Assessment of Student Performance and Progress for 2015–16*
- *Smarter Balanced Resources and Practices Comparison Crosswalk*

Site visitors were instructed to use these documents if school staff lacked familiarity with any aspect of the Smarter Balanced designated supports and accommodations. Site visitors were also instructed to make note of the areas where school staff lacked awareness or understanding of Smarter Balanced designated supports and accommodations. Site visitors offered staff hard copies of the resources to keep for personal use.

The CAASPP coordinator interview protocols focused on issues related to the process of coordinating and implementing the Smarter Balanced assessments and the provision of designated supports and accommodations. Additional questions asked respondents about feedback they might have received from teachers and students regarding the appropriateness and accessibility of designated supports and accommodations.

We developed separate protocols to observe instructional and assessment contexts. The format of these protocols drew in large part from the *Smarter Balanced Resources and Practices Comparison Crosswalk.* Site visitors completed a General Information section and Context section to provide background on elements such as grade level, content area, number of students present (along with number receiving accommodations), and number of staff present. Four focus areas were included to prompt site visitors to include notes related to *Targeted Supports/Accommodations*, *Barriers to Implementation (Teacher)*, *Barriers to Implementation (Students)*, and *Increased Student Access.* Finally, a list of instructional or assessment supports/accommodations (adapted from the *Crosswalk*) was provided for site visitors to indicate all observed instructional practices or testing supports/accommodations. Site visitors were instructed to take notes about each observed support/accommodation.

We developed two teacher interview protocols, one that focused on instructional issues and one that focused on assessment issues. Site visitors could conduct the interview prior to or after the associated observation. If the instructional interview was conducted prior to the observation, a subset of interview questions asked the teacher to describe the planned lesson. Otherwise, teachers were asked to describe what happened during the observation so that site visitors could clarify any aspects of the observation that were unclear. Both interviews included questions that targeted the similarities and differences between regularly used instructional approaches and available testing supports and accommodations.

Site visitors could use the site visit protocols in electronic or paper format, depending on what was feasible and least obtrusive in each particular setting. Completed paper protocols were entered into an electronic format for subsequent analysis.

Analysis

In addition to submitting completed observation and interview protocols, each site visitor summarized the school visits within each LEA. These summaries described differences and similarities between schools with regard to test administrator training, knowledge of students' Individualized Education Programs (IEPs), instructional approaches, observations of student testing, student engagement, integration of technology in classrooms, and any observable links between instruction and testing supports. Qualitative data from the site visit protocols and school summaries were analyzed along with data from the focus groups. All data were analyzed using NVivo 11™. The data were first coded into four main themes to reflect the four research questions. Within each theme, data were further coded into related topic areas. Finally, these topic areas were integrated into a narrative summary for each research question. This information is presented in chapter 6 of this report.

Teacher Reflection Checklist

To collect more complete data on the use of instructional practices and testing accommodations, we created a teacher reflection checklist. The purpose of this checklist was to allow teachers to document the instructional practices and classroom assessment (i.e., formative, diagnostic, interim, or benchmark testing) accommodations used with individual students over the course of the school year. These data would be combined with extant data on the accommodations offered during the summative assessment to determine the extent to which testing accommodations mirrored instructional and classroom assessment practices.

Checklist Development and Implementation

We developed the teacher reflection checklist based on the *Smarter Balanced Resources and Practices Comparison Crosswalk.* Each pedagogical practice contained in the *Crosswalk* was listed in the checklist, along with "other" options in which teachers could describe additional practices that were not included in the *Crosswalk*.

The checklist was designed so teachers would provide the student identifier (SSID) for each student, along with the student's SWD and EL status. For each listed practice, the list prompted the teacher to indicate if it was used during instruction, classroom assessment (i.e., formative, diagnostic, interim, or benchmark testing), or both, as well as an indication of the frequency of use (tried once, used occasionally, used frequently, or used daily). One checklist file included space for reporting up to three students, and teachers were instructed how to complete multiple checklists if they wanted to provide data for more than three students.

During the interview portion of the school site visit, site visitors asked participating teachers to complete the checklist. The site visitor then trained teachers on how to complete the checklist and transmit the data to HumRRO. A Universal Serial Bus (USB) flash drive containing the checklist and a document providing step-by-step instructions for completing and submitting the checklist was provided to each teacher, or in some cases the checklist and instructions were downloaded to the school POC's computer.

Teachers were given four to five weeks to complete and submit the checklists. The teacher reflection checklist is presented in Appendix D4.

Secure Data Transfer

For each teacher who agreed to complete the reflection checklist, we created an account that allowed them to access HumRRO's secure file transfer protocol (SFTP) repository. Teachers uploaded the file(s) and then notified the Study Director the file(s) had been uploaded. Completed files were then removed from the SFTP and saved to HumRRO's secure internal server. Near the end of the data collection period, which coincided with the end of the school year, we sent e-mail reminders to obtain checklists from any teachers who had not yet submitted.

Analysis

Completed checklists were received from 15 teachers who reported the use of instructional and testing accommodations for a total of 42 students. Among the students for whom data were reported, 16 (38%) were classified as SWD, 10 (24%) as EL, and 12 (29%) as both SWD and EL. No classification information was provided for 4 (10%) students in the sample.

Checklist data were first analyzed to determine the instructional practices and testing accommodations most frequently used during the school year. Table 4.5 presents a summary of the usage of instructional practices among the 42 students for whom reflection checklists were completed.

The first three columns of the table address practices and supports used for **both instruction and classroom testing**, the middle three columns address those used for **instruction only**, and the final three columns address those used for **classroom testing only**. For this study, classroom testing includes formative, diagnostic, interim, or benchmark testing. For ease of presentation, Table 4.5 focuses on the top ten selected instructional practices/supports for each type of use (instruction and classroom testing, instruction only, classroom testing only). The leftmost columns of Table 4.5 show, for example, that for 34 of the 42 students (81%) for whom checklist data were submitted, scratch paper or notepaper was used during instruction and classroom testing, and was typically used on a daily basis.

Both instructional practices and testing accommodations are determined by IEP teams; however, teachers who participated in focus groups/interviews generally reported that they are not closely involved in the process of assigning the specific summative and interim designated supports and accommodations that will be offered to students. Rather, the responsibility for uploading accommodations data into the Test Operations Management System (TOMS) generally belonged to LEA staff or the school CAASPP Coordinator. Staff in three LEAs indicated the process for entering student accommodation information was cumbersome and time consuming.

Table 4.5. Most Frequently Reported Instructional and Classroom Testing Supports/Accommodations (N=42)

Instructional Practice/ Support	*Used for both instruction and testing* N (%)	*Frequency of Use*	Instructional Practice/Support	*Used for Instruction only* N (%)	*Frequency of Use*	Instructional Practice/ Support	*Used for Testing only* N (%)	*Frequency of Use*
Scratch paper/Notepaper	34 (81%)	Daily	Math manipulatives (e.g., rulers, protractors, number lines)	29 (69%)	Frequently	Simplified test directions	6 (14%)	Occasionally
Brainstorming ideas	22 (52%)	Daily	Identify text features for understanding new vocabulary	27 (64%)	Frequently	Breaks	4 (10%)	Occasionally
Highlighter	19 (45%)	Frequently	Electronic or paper English dictionary	26 (62%)	Frequently	Magnifying devices	3 (7%)	Occasionally
Breaks	18 (43%)	Occasionally	Counting devices (e.g., blocks, tiles, chips, etc.)	25 (60%)	Occasionally	Process of elimination	3 (7%)	Occasionally
Special seating arrangements	18 (43%)	Daily	Instructional videos/movies	23 (55%)	Occasionally	Read aloud instructions	3 (7%)	Frequently
Using keyboarding to complete class assignments	16 (38%)	Frequently	Whiteboards	21 (50%)	Daily	Mark items or questions not yet answered	2 (5%)	Occasionally
Multiplication table	16 (38%)	Frequently	Automated spell-check device	20 (48%)	Frequently	Math manipulatives (e.g., rulers, protractors, number lines)	1 (2%)	Frequently
Process of elimination	15 (36%)	Frequently	Calculator (handheld or computer)	19 (45%)	Occasionally	Special seating arrangements	1 (2%)	Frequently
Read aloud instructions	15 (36%)	Daily	Electronic or paper English thesaurus	17 (41%)	Frequently	Scribes	1 (2%)	Occasionally

Notes: Instructional practices and supports adapted from the *Smarter Balanced Resources and Practices Comparison Crosswalk.* Percentages reflect the percentage of students for whom each practice/support was indicated as used for each purpose (instruction, classroom testing, or both). For this study, "classroom testing" includes formative, diagnostic, interim, or benchmark testing. Frequency of use (occasionally, frequently, daily) reflects the modal frequency category selected for each practice/support used for that purpose.

Table reads: For 34 of the 42 students (81%) "Scratch paper/Notepaper" was used during instruction and classroom testing, and teachers tended to report the frequency of use as "Daily." For 29 of the 42 students (69%) math manipulatives were used during instruction only, and teachers tended to report using them "Frequently." For 6 of the 42 students (14%) simplified test directions were used during classroom testing only, and teachers tended to report using them "Occasionally."

Interim Assessment Administration Study Survey Data

Two of the three statewide surveys about Smarter Balanced Interim Assessment Test Administration that HumRRO administered to a sample of site coordinators and test administrators in 2016, discussed in Chapter 3, included questions regarding testing accommodations and supports that are discussed in this chapter. Because LEA coordinators are more removed from actual administration of the interim assessments, they were not asked questions about accommodations and supports. More information on development and data analysis relative to the survey can be found in Chapter 3.

Table 4.6 shows the survey question numbers related to accommodations and supports on the Smarter Balanced Interim Assessments with their corresponding appendix table numbers for individual item descriptive statistics. The site coordinators and test administrators were asked to what extent the Smarter Balanced Interim Assessments incorporate supports for students with disabilities (SWD) and English learners (ELs).

Table 4.6. Interim Assessment Accommodations Survey Questions and Appendix Tables Cross-Reference

Survey	Question Numbers	Appendix C12 Tables
SITE	Q9 – Q10	62 – 63
TA	Q8 – Q9	64 – 65

As reported in Chapter 3, 107 site coordinators and 1,044 test administrators submitted survey responses and were included in the analytic sample. The test administrator respondents represented 376 different LEAs and 981 different schools. The characteristics of our analytic sample are not broadly representative of the state as a whole and limit our ability to generalize our findings. The survey findings we report must therefore be considered within the context of our sample.

School site coordinators generally agreed the accommodations and designated supports of the interim assessments contributed to positive experiences for SWDs. Their level of agreement to the statements that the interim assessment accommodations and designated supports improved access to the test (61%–67% strongly agreed or agreed) and that the accommodations and designated supports functioned properly during testing (61%–64% strongly agreed or agreed) (see Figure 4.1).

Test administrators generally agreed the accommodations and accessibility features of the interim assessments contributed to positive experiences for SWDs (see Figure 4.2); however, their responses were not as favorable as those of the school site coordinators. For example, less than half of test administrators agreed that for SWDs, the accommodations and designated supports (a) improved access to the test, (b) are consistent with instructional practices used in the classroom, and (c) functioned properly during operational testing. Additionally, a large percentage of test administrators reported not knowing whether the Smarter Balanced Interim Assessments resulted in positive experiences for SWDs (as indicated by “I don’t know”).

In ratings made by test administrators for EL students, the highest level of agreement (51%) was with the statement that accessibility features functioned properly. Similar to SWDs, the majority of test administrators expressed disagreement with statements about improved access to the test, consistency of testing features with classroom practices, students having a familiar, comfortable and problem-free testing experience, and students being able to demonstrate knowledge and skills on the interim assessments. (See Figure 4.2).

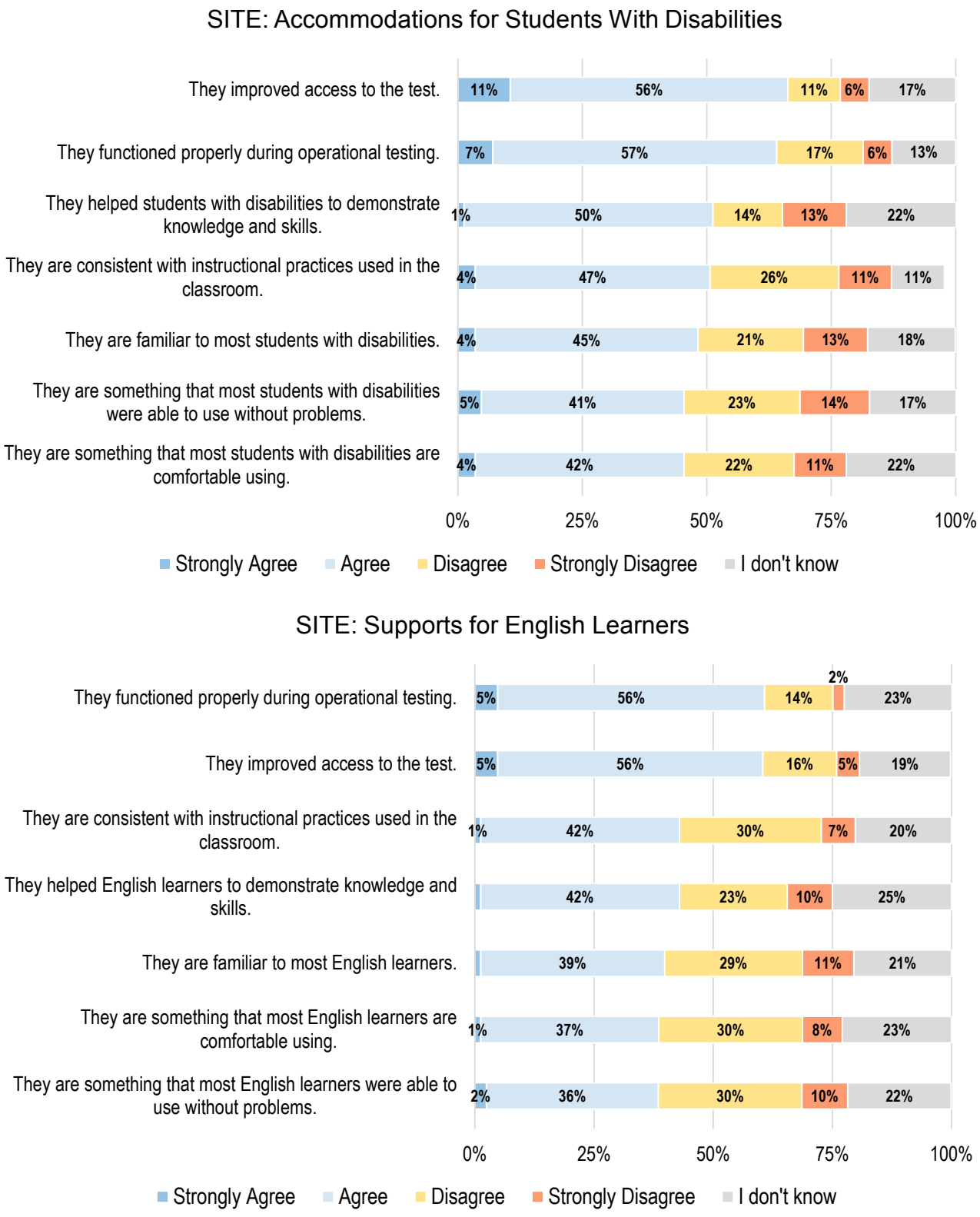
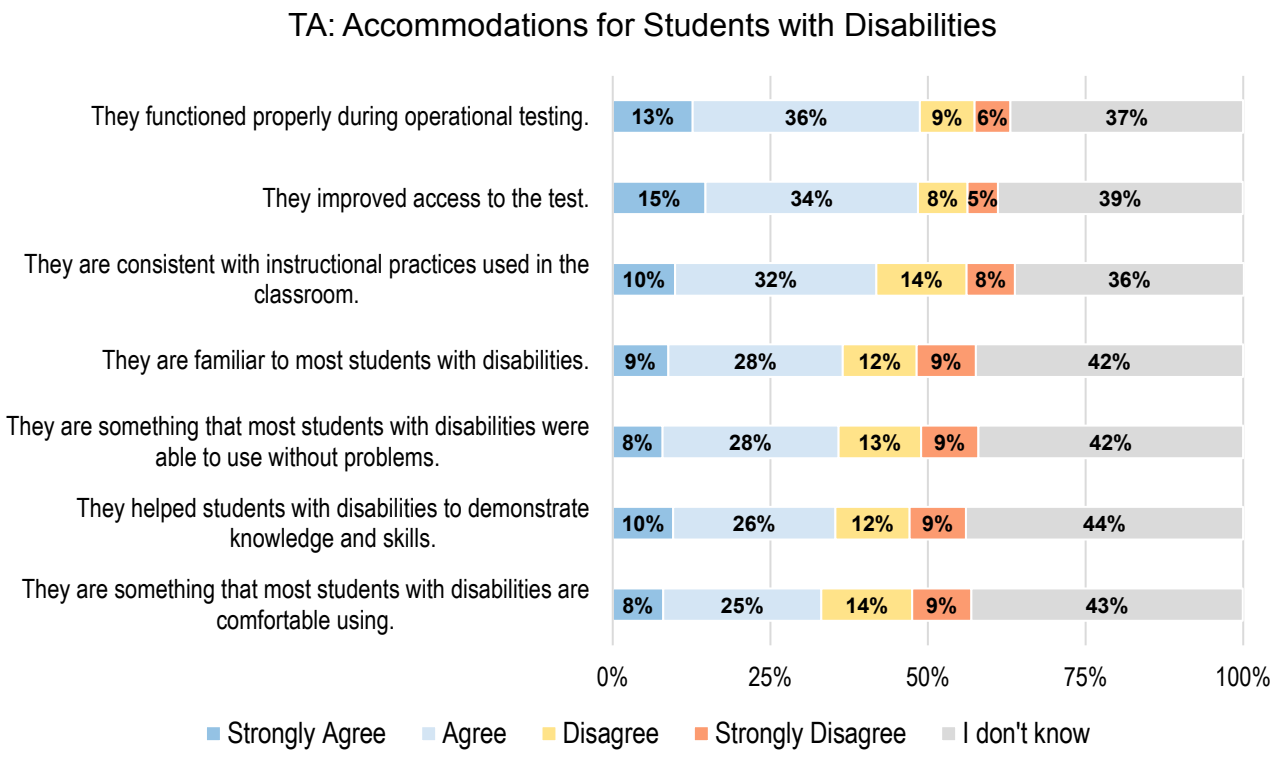


Figure Note: Path 1 analytic sample for each survey: N_{LEA} = 285; N_{SITE} = 90; N_{TA} = 863

Figure 4.1. Accommodations for students with disabilities and supports for English learners (SITE survey).



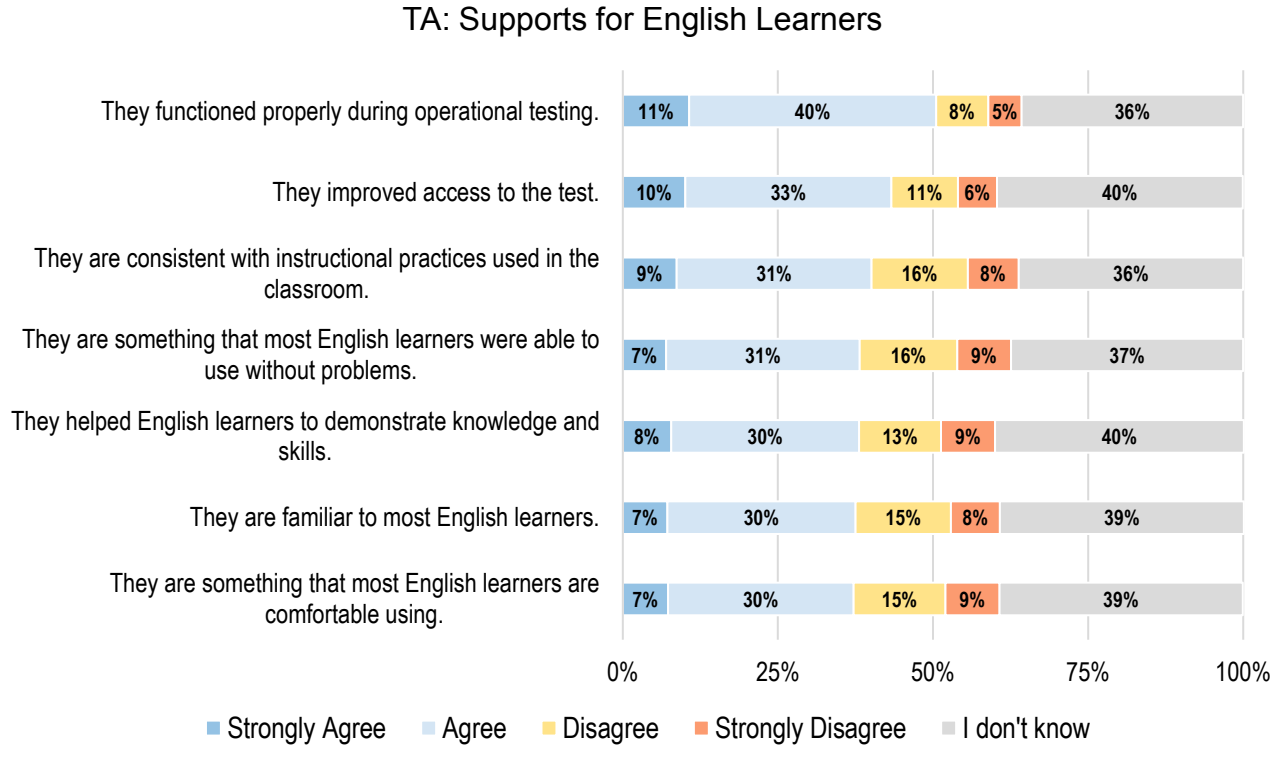


Figure Note: Path 1 analytic sample for each survey: NLEA = 285; NSITE = 90; NTA = 863

Figure 4.2. Accommodations for students with disabilities and supports for English learners (TA survey).

Data on Summative Assessment Accommodations and Supports

The CDE provided HumRRO with a file containing data on the embedded and non-embedded accommodations and designated supports offered to students during the 2015–16 Smarter Balanced Summative Assessments in ELA and mathematics. Student names had been removed, but the Statewide Student Identifier (SSID) was retained to allow us to merge data from the completed Teacher Checklists. Table 4.7 presents the number of students that were included in the analysis of overall provision of summative assessment accommodations and designated supports.

Table 4.7. Percentage of Students Taking 2015–16 Smarter Balanced Assessments Identified as SWD or EL

Total Number of Students Tested	Students with Disabilities	English Learners with Disabilities	English Learners
English Language Arts/Literacy (n=3,339,029)	7.42% (n=247,819)	3.39% (n=113,122)	15.05% (n=502,588)
Mathematics (n=3,339,002)	7.42% (n=247,800)	3.39% (n=113,117)	15.05% (n=502,595)

Analysis

The data file contained several variables, each representing a different accommodation or designated support. Using *Matrix One: Universal Tools, Designated Supports, and Accommodations for the California Assessment of Student Performance and Progress for 2015–16* as a guideline, HumRRO derived new variables indicating if an accommodation or designated support was offered during any portion of the test (e.g., Performance Task (PT) and non-PT), and the total number of embedded accommodations, embedded designated supports, non-embedded accommodations, and non-embedded designated supports offered to each student. These variables were created separately for ELA, taking into account that accessibility features may be classified differently depending on the content type and test section (PT or non-PT).

Table 4.8 presents the percentage of students with disabilities, English learners with disabilities, and English learners offered at least one summative assessment accommodation, designated support, or universal tool, by content area. The majority of students with disabilities, regardless of English Learner status, were offered at least one accommodation or designated support. The large majority also had access to embedded universal tools, unless access was turned off as part of a designated support[38]. Accommodations are offered only to students with an IEP or Section 504 plan. Nearly all English learners were offered access to universal tools during the ELA assessment and were offered access to the universal tools and/or at least one designated support during the math assessment.

The last row in Table 4.8 presents the comparable percentages for students not identified as a student with a disability, English learner with a disability, or English learner. These non-identified students were less likely than students with disabilities to have been offered designated support in ELA or an accommodation in math. The percentages of students receiving math designated supports is high across all the subgroups due to the fact that nearly all students were offered the *translations (glossary)* designated support (available for math only).

[38] By default, universal tools are available to all students. If needed, universal tools may be turned off. Turning off universal tools is a designated support.

Table 4.8. Percentage of Students Offered Accommodations, Designated Supports, and Universal Tools*

	ELA			Math		
	Accommodations	Designated Supports	Universal Tools	Accommodations	Designated Supports	Universal Tools
Students with Disabilities	19.72 (n=48,871)	51.69 (n=128,097)	92.83 (n=230,056)	23.67 (n=58,661)	99.97 (n=247,729)	92.46 (n=229,108)
English Learners with Disabilities	26.80 (n=30,312)	61.25 (n=69,287)	95.39 (n=107,905)	30.09 (n=34,035)	99.85 (n=112,948)	95.21 (n=107,704)
English Learners	0.08 (n=419)	20.01 (n=100,549)	94.95 (n=477,211)	0.08 (n=421)	99.77 (n=501,464)	96.94 (n=487,214)
Non- identified	0.08 (n=2,097)	7.37 (n=182,467)	96.79 (n=2,395,926)	0.09 (n=2,151)	99.90 (n=2,473,104)	96.74 (n=2,394,822)

* Includes both embedded and non-embedded accommodations and designated supports.

Data on summative assessment accommodations and designated supports were merged with Teacher Checklist data using the SSID provided by the teachers completing the checklist. The SSIDs yielded 34 out of 42 successful merges[39] (see Table 4.9). Next, a crosswalk was created between the checklist variables and the summative assessment accessibility features to identify the corresponding instructional practices. For each pair of instructional practices and accommodations/designated supports/universal tools, a new variable was then created to indicate that the student was offered both.

Data were analyzed via a combination of frequency counts and measures of central tendency and variability.

Table 4.9 Students with Matched Checklist and Summative Assessment Data

	Students with Disabilities	English Learners with Disabilities	English Learners
Matched Students (n=34)	15 (44%)	11 (32%)	3 (9%)

Among the small sample of students for whom data on both instructional practices and summative accommodations/designated supports were available, the majority (28 of 34 for ELA; 32 of 34 for math) were offered summative accommodations, designated supports, or universal tools that corresponded with at least half of the instructional practices reported by teachers. Table 4.10 presents the most frequently offered instructional practices and the percentage of students who were offered the matching summative accommodation, designated support, or universal tool. There were several commonly reported instructional practices for which students were not offered the corresponding summative accommodation, designated support, or universal tool that is allowable within the Smarter Balanced assessment system. These summative accommodations/designated supports included closed captioning, multiplication table, abacus, and streamlining. We are unable to determine from the available data if these features were not offered based on a decision made by staff responsible for assigning

[39] Five teachers provided SSIDs that did not match any SSID in the summative assessment dataset. We followed up with these teachers but were unable to get correct SSIDs for these 8 students.

accommodations/designated supports within the TOMS system, or if this reflects a lack of staff familiarity with the available accommodations/designated supports.

Table 4.10. Match Between Instructional Practices and Summative Assessment Accommodations, Designated Supports, and Universal Tools (N=34)

Most Frequently Used Instructional Practices Among Matched Sample	% Receiving Matching Accommodation/Designated Support/Universal Tool	Matched Summative Assessment Feature(s)
Scratch paper/Notepaper (n= 32)	97	Universal Tools On
Math manipulatives (n= 29)	97	Universal Tools On
Brainstorming ideas (n= 27)	93	Universal Tools On
Electronic or paper English dictionary (n=27)	93	Translation (glossaries)
Identify text features for understanding new vocabulary (n= 26)	96	Universal Tools On
Highlighter (n= 26)	100	Universal Tools On
Read aloud instructions (n=26)	54	Text-to-Speech; Read Aloud items/passages
Proofreading (n=25)	96	Universal Tools On
Using keyboarding to complete class assignments (n= 25)	92	Universal Tools On
Instructional videos/movies (n=25)	0	Closed Captioning
Multiplication table (n=25)	0	Multiplication Table
Special seating arrangements (n=23)	26	Separate Setting; Noise Buffers
Breaks (n=22)	100	Universal Tools On
Counting devices (n=22)	0	Abacus
Automated spell-check device (n=22)	95	Universal Tools On
Calculator (n=22)	95	Universal Tools On
Process of elimination (20)	95	Universal Tools On
Writing down connections n=18)	89	Universal Tools On
Simplified test directions (n= 17)	0	Streamlining

One way that this issue could manifest would be individual students being offered a large number of accommodations and designated supports in addition to the universal tools provided, even if these were not typically provided in their classroom. School or district staff might think that “more is better” when it comes to accessibility features, and select as many as possible. To explore this possibility, we looked at the total numbers of accommodations and designated supports offered to all students with disabilities, English learners with disabilities, and English learners who took the 2015–16 Smarter Balanced ELA and mathematics assessments. Table 4.11 shows that on average, students were typically offered one or two accommodations or supports (as well as access to universal tools), and that there was little variability among individual students. Only around 2 percent to 5 percent (ELA and math, respectively) of all

students with disabilities and/or English learners were offered 5 or more accommodations/designated supports. In subsequent analysis we identified no LEAs in which students with disabilities, English learners with disabilities, or English learners were systematically offered large numbers of accessibility features

Table 4.11. Average Number of Accommodations and Designated Supports Offered to Students Statewide*

	ELA				Math			
	Min	Max	Median	Mean (SD)	Min	Max	Median	Mean (SD)
English Proficient Students with Disabilities	0	17	1	1.13 (1.41)	0	18	2	2.26 (1.65)
English Learners with Disabilities	0	16	1	1.48 (1.54)	0	17	2	2.64 (1.74)
English Learners	0	12	0	0.27 (0.67)	0	14	1	1.20 (0.60)

* Includes both embedded and non-embedded accommodations and designated supports.

This page is intentionally blank.

CHAPTER 5: UPDATE TO THE CAASPP INDEPENDENT EVALUATION STUDY PLAN

Michele Hardoin

The State Board of Education (SBE) approved the California Assessment of Student Performance and Progress (CAASPP) Independent Evaluation Study Plan on September 2, 2015, as noted in Chapter 1 of this report and as fully described in the first annual evaluation report (Hardoin, Becker, Wise, 2015). HumRRO's contract with the California Department of Education (CDE) calls for the three-year plan to be reviewed and updated annually, based on findings to date and on any emerging policy issues that were not fully anticipated in the original design. This chapter presents factors considered during the 2016 review of the original plan and revisions made to the plan to align it with the current development timelines for two new components of the CAASPP System, the Next Generation Science Standards (NGSS) science assessments and the California Alternate Assessments (CAAs) in English language arts (ELA) and mathematics.

Original Independent Evaluation Study Plan

When the Independent Evaluation Study Plan was developed, studies that met the criteria for inclusion were prioritized in consideration of several factors, including which areas within the CAASPP System were currently of greatest concern, the degree of formative value of study findings and recommendations, and the readiness or stability of the aspect of the assessment to be studied. Although a myriad of studies have the potential to provide information about and suggest improvements for the CAASPP System, limited resources of the CDE constrain the evaluation to conducting two research studies per full school year (2015–16 and 2016–17) and one study in 2016–17, to be completed by the December 31, 2017 contract end date.

HumRRO's original Independent Evaluation Study Plan included the following schedule for conducting five independent evaluation studies:

- During the 2015–16 school year:
  1. Interim Assessment Test Administration Study
  2. Access to Designated Supports and Accommodations Study
- During the 2016–17 school year:
  3. Review of Scoring Processes Study
  4. Utility of Score Reporting Study
- During the 2017–18 school year:
  5. Item Alignment and Quality of Science Assessments Study

The final study, an alignment study, was planned in anticipation of the new California NGSS science assessments being pilot tested during 2015–16. As originally scheduled, this alignment study would have been conducted prior to the implementation of the operational science assessments. It would address research questions such as whether each pilot test item in the computer-based science assessments measured targeted content clearly, if items were free from irrelevant requirements, and if the items as a whole appropriately covered the content standards.

Responding to Revised Timelines for New CAASPP Assessments

In early December 2015, CDE staff and HumRRO staff met in person to review progress on the 2015–16 independent evaluation studies and to review and discuss possible changes in priority or context for the studies scheduled to be conducted during the 2016–17 and 2017–18 school years. The CDE confirmed the plan to move forward with the two 2016–17 studies, the Review of Scoring Processes Study and the Utility of Reporting Study. However, HumRRO learned that delays in development and administration of the science assessments would no longer support the planned schedule for conducting the independent alignment study. The new science assessments would not be pilot tested until spring of 2017, with field testing in 2018 and operational assessments implemented in 2019.

We identified a logical replacement for the science alignment study, an alignment study of another new CAASPP component, the California Alternate Assessments (CAAs). California had been working on development of an alternate assessment for English language arts/literacy (ELA) and mathematics with the National Center and State Collaborative (NCSC), to replace the California Alternate Performance Assessment (CAPA) that ended in 2014. California anticipated using the NCSC assessment which had been expected to be operational by 2015. However, issues with the small student count in the NCSC field test and California's demand for a census field test led to California's decision to develop its own alternate assessments with its testing contractor, the Educational Testing Service (ETS). The new CAAs are planned to be pilot tested in 2016, with operational assessments implemented in 2017, and they differ in format and testing methods from the former alternate assessments. Although detailed test specifications for the new CAAs were not yet available in December 2015, the CDE believed an independent alignment review of these tests would be beneficial and would meet the criteria for independent evaluation studies.

With this new direction from the CDE, HumRRO developed a redesigned alignment study plan for reviewing the 2017 operational forms of the CAAs for ELA and mathematics for grades three through eight and grade eleven. We removed the science assessments alignment review and incorporated the new alignment study into a revised Independent Evaluation Study Plan. The updated three-year plan was presented to the CDE and its Technical Advisory Group (TAG) at the spring 2016 TAG meeting. Discussion of the updated three-year plan and evaluation studies included participation of TAG members, CDE staff from the Assessment Development and Administration Division (ADAD) and the Accountability, Measurement, and Reporting Division (AMARD); and ETS staff.

Summary of 2016 Updates to the Independent Evaluation Studies

The 2016 revisions to the Independent Evaluation Study Plan respond to the changing timelines of the development and administration of the CAASPP science assessments and the CAAs. Following are updated descriptions of the three remaining research studies to be conducted in the three-year plan. Each description includes the current planned timing of each study, the theoretical rationale for each study, and the research questions to be addressed. Specific details about the methods to be used in each study were developed in consultation with the CDE and the TAG and will be included in future interim annual evaluation reports.

Review of Scoring Processes Study (2016–17)

The Smarter Balanced assessments consist of closed-ended items that can be machine scored immediately and test questions with open-ended responses requiring trained human scorers.

Open-ended response scoring processes are susceptible to challenges that closed-ended item scoring is not. For example, human scorers must be trained on general scoring processes and how to score consistently within an overall rubric, as well as on the specific demands of individual test items. Scorers must be monitored using calibration and validation sets, to identify and correct scoring issues with further training or other actions.

HumRRO plans a two-part study that can be useful to ensuring that scoring processes are reliable and lead to valid inferences of CAASPP results. First, HumRRO will conduct a process evaluation of human scoring for the **summative** Smarter Balanced assessments. An independent review of the scoring process from scorer training through monitoring and reporting procedures and an in-depth review of scoring materials for a sample of items are vital evidence to collect to support valid score inferences. Second, HumRRO will conduct a collaborative study with a sample of LEAs from the LEARN to investigate the usage and quality of **interim** assessment scoring training and scorer materials. Research questions for this study include:

- Are the training and monitoring procedures for hand scoring Smarter Balanced Summative Assessment responses of high quality?
- Are the procedures and training for hand scoring Smarter Balanced Interim Assessments responses sufficient in breadth, depth, and clarity to support effective local scoring?

Study methods will include reviewing documents of scoring procedures, observing scorer training (in person or via Webinar), and surveying scorers for the interim assessments.

Utility of Score Reporting (2016–17)

To achieve the goals for the CAASPP System, the score reports for the Smarter Balanced summative and interim assessments must communicate accurate and useful information for the various stakeholders, but especially for teachers and leaders at schools and districts. Appropriate interpretations of test results will have positive consequences if, for example, they are used to identify and implement appropriate instructional adjustments or highlight needed areas of professional development. California employs a variety of methods for reporting scores, including individual student score reports, and would benefit by reviewing the system's functionality among its stakeholders who must use the system to perform their jobs.

HumRRO plans a two-part study to explore the use and perceived utility of (a) reports for Smarter Balanced **Summative** Assessments and (b) reports for Smarter Balanced **Interim** Assessments. The study will focus on uses by teachers and by school and LEA leaders. Research questions include:

- Does the score reporting system for Smarter Balanced assessments provide LEA and school stakeholders with actionable information that supports intended uses?
- Are there differences in the use and utility of the reporting system by grade level, subject, student subgroup, region, or availability of technical support?
- Is the score reporting system accessible and user-friendly?

Study methods will include focus groups with key stakeholders from a sample of LEAs from the LEARN and other LEAs, as needed, and surveying LEA and school educators.

California Alternate Assessments (CAA) Alignment Study (July–December, 2017)

Potential inferences made from assessment scores or results (e.g., indications of students' overall readiness or proficiency or an individual student's areas of strength or weakness within a content domain) are built on assumptions about the qualities of the assessment. HumRRO plans to conduct an alignment study after the 2017 operational alternative assessments to inform aspects of the assessments that could lead to the improved validity of test score interpretations. The format and testing methods of the new alternate assessments are changed sufficiently that an alignment study is necessary to support the validity of the CAA scores. CAA is a tiered adaptive assessment, which presents challenges for investigating alignment, challenges that must be overcome to generate accurate interpretable alignment results.

HumRRO will carry out an argument-based logical reasoning approach, similar to a validity argument. We present five claims that must be true for the CAA to be adequately aligned to content standards and then will collect evidence to support those claims. If the claims are all supported, the overall alignment argument is supported. If there are areas where the evidence is weak or the claims are not supported, these represent threats to the alignment argument, which is also a threat to the overall test validity argument. The claims are as follows:

1. The California Alternate Assessment Standards (CAA Standards)[40] are highly related to the Common Core State Standards (CCSS).
2. The California Alternate Assessment test items represent the intended content.
3. The California Alternate Assessment test items are of similar cognitive complexity to the CAA Standards.
4. The adaptive staging algorithm of the California Alternate Assessments allows for improved access to the assessment.
5. The California Alternate Assessments are sufficiently reliable for reporting.

Study methods will include subject matter panels (California teachers and special education experts) convened to review items and provide judgments about skills required to answer the items.

[40] The CDE refers to the CAA Standards as Connectors, Prioritized Common Core Connectors (CCCs), and Essential Understandings.

CHAPTER 6: FINDINGS AND RECOMMENDATIONS

As described in Chapter 1, the three-year Independent Evaluation Study Plan for the California Assessment of Student Performance and Progress (CAASPP) was approved by the State Board of Education in September 2015. The evaluation is required under California *Education Code* (EC) Section 60649, and its purpose is to provide objective technical advice and consultation on activities supporting the continuous improvement of the CAASPP System developed and administered pursuant to EC Section 60640.

This chapter summarizes the findings and presents recommendations stemming from the following key evaluation activities we conducted during the 2015–16 school year:

- Smarter Balanced Interim Assessment Administration Study
- Access to Designated Supports and Accommodations Study
- Observations of CAASPP System training provided to California teachers, administrators, and district staff

The two research studies were planned and implemented with guidance from CDE staff and the CAASPP Technical Advisory Group (TAG). For a complete description of the methods, participants, analysis, and contextual details of each 2015–16 study, please refer to Chapter 3: Smarter Balanced Interim Assessment Administration Study and Chapter 4: Access to Designated Supports and Accommodations Study.

For each of the studies conducted during the 2015–16 school year, we present findings organized around the study's research questions and then offer our recommendations for improving the quality and effectiveness of the CAASPP System based on the study's findings. We conclude the chapter with recommendations based on our observations of CAASPP training sessions and a few comments on the progress made by the CDE in implementing the 2015–16 assessments.

Results from Smarter Balanced Interim Assessment Administration Study (Chapter 3)

The Smarter Balanced Interim Assessment Administration Study was designed to explore: (a) how LEAs decided whether and which interim assessments to administer in 2015–16, (b) how smoothly the administration and scoring of these assessments went, and (c) the extent to which the results may have had an impact on instruction. LEAs have the authority to decide whether to use interim assessments provided as part of the Smarter Balanced Assessment System, which include Interim Comprehensive Assessments and Interim Assessment Blocks. According to the CDE's *Smarter Balanced Interim Assessments* Web site,[41] the interim assessments are specifically intended to provide:

- meaningful information for gauging student progress throughout the year toward mastery of the skills measured by the summative assessments
- assessments of specific Common Core State Standards (CCSS), which can be used at strategic points during the school year

[41] http://www.cde.ca.gov/ta/tg/sa/sbacinterimassess.asp

The overall objective of this study was to understand and analyze local experiences with the Smarter Balanced Interim Assessments via a small set of focus groups and online surveys. The survey sample was designed to reflect the diversity of LEAs throughout the state, but *was not* intended to provide sufficient statistical representation of all LEAs to support precise statistical estimates of frequencies and proportions for the state as a whole. Due to local flexibility in how the interim assessments are administered, CDE and HumRRO deemed it prudent to begin with a more qualitative study of local educators' experiences and perceptions with these CAASPP assessments. Rather than a strictly representative sample, we intentionally included LEAs (in all three surveys) estimated to be high interim assessment users based on rough estimates of 2015–16 interim assessment administration counts through March 31, 2016 provided by ETS.

The study gathered data from the Local Educational Agency Research Network (LEARN[42]) through a combination of nine telephonic focus groups and interviews and 11 in-person school site visit interviews. These data were primarily used to inform the creation of three statewide surveys administered online to LEA CAASPP coordinators, school site CAASPP coordinators,[43] and Smarter Balanced Interim Assessment test administrators. A full description of the development of data collection instruments, training of interviewers, response rates, demographics of respondents, and other study details and results can be found in Chapter 3.

The initial intent of the study was to explore how use of the Smarter Balanced Interim Assessments has impacted instructional and student-level decisions about learning. We had indications from the focus group data that some LEAs were not using the interim assessments primarily for this purpose. Despite this challenge, results from the survey do inform the original research questions; however, they do not provide rich descriptions of interim assessments informing instructional practice. What we have is a broad picture of a range of experiences from each of the three stakeholder groups surveyed, based on our limited response rates, and the respondents' perceptions of notable challenges and suggestions for improvements. The implementation of a comprehensive and newly introduced assessment system with a multitude of resources and tools are bound to be met with varying levels of experiences and challenges. The CDE and Smarter Balanced have already addressed some of these "growing pains" during the 2015–16 school year, such as switching to a common login user ID for systems that previously permitted different user IDs and providing clarification about options for administering the interim assessments in a non-standardized manner. As the Smarter Balanced Interim Assessment system continues to develop and improve, our findings provide additional context for understanding the transition from initial interim assessment use in winter 2014–15 to full school year implementation in 2015–16.

Specific results for each research question follow. Results consolidate our findings from the three statewide online surveys. Although we do not provide summaries of our focus group and interview findings, whose purpose was primarily to inform survey development, we do apply information gleaned from our qualitative data collection to help understand possible reasons for certain survey response frequencies. The percentages reported represent survey results, rounded to whole number percentages from a relatively large and diverse set of survey respondents. The respondent counts for the surveys were as follows: LEA CAASPP coordinators (N=367), school site CAASPP coordinators (N=107), and Smarter Balanced Interim Assessment test administrators (N=1,044). Note that the first question in each survey established respondent eligibility (e.g., "Did you administer a Smarter Balanced Interim

[42] See Chapter 2 for a full description of the LEARN.
[43] Hereafter, LEA CAASPP coordinators will be referred to as "LEA coordinators" and school site CAASPP coordinators will be referred to as "site coordinators."

Assessment this school year?"), and respondents who did not administer any Smarter Balanced Interim Assessments were routed to a short series of questions. Overall, more than three-fourths of the respondents to each survey reported administering at least one Smarter Balanced Interim Assessment during 2015–16 (78% of LEA coordinators, 84% of site coordinators, and 83% of test administrators). Percentages reported in conjunction with subsequent questions represent respondents in our sample who administered one or more of the interim assessments. Below we summarize evidence that supports key findings for each research question and provide some background information to help contextualize the perceptions of the educators in our sample.

A few caveats are important to consider when interpreting the results and recommendations from this study. First, staff from LEAs and administrators and teachers from schools who did not participate in any phase of the study might have provided a different view of the Smarter Balanced Interim Assessments. Second, our relatively low online survey response rates (about 22% for LEA coordinators, about 10% for site coordinators, and about 9% for test administrators), limit our ability to generalize across the state. Third, we expanded our original study design to include perspectives of LEA and site coordinators in addition to test administrators. However, we did not select our target sample for the site coordinators and test coordinators as a matched sample to ensure all LEAs in our sample were represented by each of the other stakeholder groups. Therefore, we are limited in determining to what extent LEAs, schools, and test coordinators agree or disagree based on their experiences.

Interim Assessment Study Research Question 1. How are decisions made about whether and how interim assessments are used?

To help answer this first research question, our surveys included questions about factors that impacted LEA decisions to use the interim assessments. Because of our sample limitations, generalizations *cannot* be made as to how widespread these policies might be, but they provide an indication of whether LEAs serving certain populations use the interim assessments.

IA Study Key Finding 1.1. In our sample,[44] the primary factor that contributed to LEA and school decisions to require or highly encourage use of the Smarter Balanced Interim Assessments was the value of familiarizing students with the testing systems–including universal tools, designated supports, and accommodations– and the types of questions that are included in the Smarter Balanced Summative Assessments.

Supporting Evidence and Additional Information:

- At least 78% of the LEA coordinators and 75% of site coordinators who responded to our survey and reported requiring or highly encouraging interim assessment usage cited these factors.
- The second most reported primary factor for choosing to use the interim assessments was to inform classroom instruction (59% of LEA coordinators and 64% of site coordinators).

[44] The survey sample of potential respondents included 1,599 LEA CAASPP coordinators (all remaining cases after removing duplicates from the roster of 1,944 cases); 1,006 school site CAASPP coordinators (random selection of every ninth school after removing duplicates from the roster of 9,040 cases); and 12,751 test administrators (random selection of every 15th case after removing duplicates from the roster of 191,265 cases). See Chapter 3 for additional sampling details.

The most commonly cited uses of the interim assessments that contributed to requiring or highly encouraging their use was familiarizing students and teachers with testing/testing systems (LEA = 93%, 213 of 229; SITE = 85%, 72 of 84[45]), familiarizing students and teachers with item types (LEA = 83%, 190 of 229; SITE = 83%, 70 of 84), and providing students practice with universal tools, designated supports, and accommodations (LEA = 78%, 179 of 229; SITE = 75%, 63 of 84). Over half of LEA coordinators (59%, 136 of 229) and school site coordinators (64%, 54 of 84) reported "to inform classroom instruction" as a significant reason for using the interim assessments. Smaller percentages of coordinators reported their primary uses as informing subject matter professional development offerings or determining student placement.

IA Study Key Finding 1.2. In our sample, the hand scoring requirement of some Smarter Balanced Interim Assessments was not a major factor in deciding whether to administer those assessments.

Supporting Evidence:

- 75% of test administrators who responded to our survey reported hand scoring was *not* a factor in assessment selection.

The cost and complexity of scoring open-ended responses did not appear to widely influence the decision about which Smarter Balanced Interim Assessments to administer. Even though the hand scoring component does require extra time, training, and labor to complete, only 16 percent of test administrators reported they specifically administered interim assessments that *did not* require hand scoring. It is unclear to what extent LEA-wide policies of encouraging or requiring IABs that did not require hand scoring impacted this finding. The remaining 8 percent of test administrators reported they intentionally selected interim assessments that *required* hand scoring.

IA Study Key Finding 1.3. In our sample, a significant percentage of LEAs required their schools to administer Smarter Balanced Interim Assessments during 2015–16, but others left the decision to their schools.

Supporting Evidence:

- 47% of LEA coordinators, 54% of school coordinators, and 82% of test administrators who responded to our survey reported their LEA required some use of Smarter Balanced Interim Assessments.

Reported policies about the required use of Smarter Balanced Interim Assessments were mixed. Of those in our sample administering at least one interim assessment, 134 of 285 LEA coordinators reported that their LEA required the Smarter Balanced Interim Assessments, as did 49 of 90 site coordinators, and 705 of 863 test administrators. The remaining respondents reported the interim assessments were either voluntary but highly encouraged or voluntary with no specific encouragement.

Our more detailed survey data analysis found that, of the 54 LEA coordinators from charter schools, 34 reported the Smarter Balanced Interim Assessments were required compared to 100 of the 227 LEA coordinators from districts. Of the 166 LEA coordinators from LEAs with a

[45] LEA = respondents to LEA CAASPP Coordinator survey, SITE = respondents to school site coordinator survey, and TA = respondents to test administrator survey.

high EL population, 89 reported the interim assessments were a requirement compared to 40 of the 105 LEA coordinators from LEAs with a low EL population.

IA Study Key Finding 1.4. In our sample of LEAs, discrepancies in responses within an LEA as to whether the Smarter Balanced Interim Assessments were required suggest LEA policies were not clear to some test administrators.

Supporting Evidence:

- We matched LEA coordinator and test administrator responses for 139 LEAs. For 67 LEAs, 84% of test administrators who responded to our survey agreed with LEA coordinators that the LEA policy required some use of Smarter Balanced Interim Assessments. For the other 72 LEAs, only 25% of test administrators agreed with LEA coordinators that use of Smarter Balanced Interim Assessments was *not* a requirement.

There were instances of a discrepancy between test administrator and LEA-level reports of interim assessment policy from within the same LEA. This suggests that either (a) LEA coordinators and test administrators have differing definitions of what a requirement is versus something that is highly encouraged, (b) the school might have a requirement that test administrators consider an LEA-wide policy, or (c) there is miscommunication or misunderstanding of the policies. Our more detailed survey data analysis found that, within our sample, the reported policy discrepancies do not appear to be related to the size of the LEA.

This finding suggests that differentiating the policy for the Smarter Balanced Interim Assessments from other third-party or district formative assessments could be beneficial in making test administrators, who are often teachers, more aware—not only of the collection of assessment tools available to them, but also of the myriad ways in which student performance can be measured. Making the expectations clear to teachers is key in how they approach using the assessments. Additionally, if teachers are potentially mixing up the Smarter Balanced Interim Assessments with third-party or district interim assessments, that could also influence their perceptions of how well the Smarter Balanced Interim Assessment system is working.

Interim Assessment Study Research Question 2. What are detectable patterns in the types of interim assessments used (e.g., interim comprehensive assessments versus interim assessment blocks)?

We obtained empirical data from ETS to help answer research question 2. Our survey responses provide information on which assessments are required by the LEAs and about general usage for our specific study sample. Further, this study focused on identifying factors that influenced the experiences and usage of the Smarter Balanced Interim Assessments. We learned from our focus groups that some LEAs who administered the Smarter Balanced Interim Assessments also used third-party assessment systems, and that for some LEAs the use of third-party interim assessments is the reason the LEA does not use the Smarter Balanced Interim Assessments. Although CDE and Smarter Balanced do not advocate or mandate the use of one particular interim assessment system, we included questions about third-party interim assessment systems on our surveys to further explore this situation.

IA Study Key Finding 2.1. Statewide, large numbers of schools and LEAs administered Smarter Balanced Interim Assessments during 2015–16.

Supporting Evidence:

- ETS interim assessment usage data, through March 2016, show that 4,321,412 interim assessments were administered. Based on HumRRO's matching of interim assessment data to state-mandated summative assessment data, a total of 6,178 schools out of 9,212 (about 67%) indicated using at least one interim assessment compared to only 3,034 that did not (33%).

ETS interim assessment usage data, through March 2016, show that larger numbers of schools administered IABs (5,774) compared to ICAs (3,703) and the majority of schools who administered IABs also administered ICAs (3,373 administered both). In our sample, the mathematics and ELA/literacy IABs were the most commonly reported Smarter Balanced Interim Assessments required or highly encouraged by LEA and school site coordinators. Perhaps, as was discussed in our focus groups, the higher rates of encouragement of the IABs are due to the shorter administration time and, for some IABs, the lack of hand scoring requirement. The variations in LEA coordinator, site coordinator, and test administrator reports may be partially explained by differing school- and LEA-level policies and by the fact that the responses from the three stakeholder groups did not represent the same LEAs, although there were overlapping responses among groups for some LEAs.

IA Study Key Finding 2.2. Although this study did not specifically explore the role or quality of third-party assessment systems, many respondents in our sample indicated that third-party assessment systems were used in 2015–16 in combination with Smarter Balanced Interim Assessments.

Supporting Evidence:

- 71% of LEA coordinators who responded to our survey and administered Smarter Balanced Interim Assessments reported use of some other third-party assessment system as well.
- 76% of LEA coordinators who responded to our survey but did *not* administer the Smarter Balanced Interim Assessments reported use of some other third-party assessment system.

Of the 273 LEA coordinators who administered Smarter Balanced Interim Assessments, 194 of them reported also using some other third-party assessment system. Of the 81 LEA coordinators who indicated *not* administering a Smarter Balanced Interim Assessment, 24 percent reported that the reason was because their current third-party interim assessment platform (e.g., Illuminate DnA, Renaissance Place) provides more benefits than the Smarter Balanced system.

In our sample, there was no indication that a specific third-party assessment is the preferred platform. There is some evidence to suggest that larger LEAs might utilize the Smarter Balanced Interim Assessments more than smaller LEAs and that the use of third-party systems influenced whether LEAs administered the Smarter Balanced Interim Assessments. From our more detailed analysis, of the 52 small LEAs in our sample, 54 percent indicated *not* using a third-party system compared to 21 percent of large LEAs. This study, however, did not explore specifics about why some coordinators in our sample perceived their current third-party system as more beneficial than the Smarter Balanced Interim Assessments.

Because the use of Smarter Balanced Interim Assessments is decided by LEAs, who have flexibility to use a variety of tools and resources to help promote student performance and progress, we note that the role of third-party assessment systems appears to have some uncertain influence on how the Smarter Balanced Interim Assessments are being used. As this study did not focus on providing an in-depth exploration of the third-party assessments, we cannot provide information regarding the quality or features of these systems compared to the Smarter Balanced Interim Assessments. Our findings suggest LEAs are still evaluating the relative strengths and weaknesses of the Smarter Balanced Interim Assessments and third-party assessments.

Further study could provide insight into how LEAs' considerations of multiple data sources of interim assessments, specifically third-party interim assessment systems used in addition to the Smarter Balanced Interim Assessments, enhance student learning. The study could investigate (a) how LEAs use their third-party system, (b) the benefits and drawbacks of that system relative to the Smarter Balanced Interim Assessment System, and (c) potential ways to combine data across different assessment systems or to develop other integrative features to facilitate working with data from separate sources. The results of such a study might also help inform LEAs during their selection and design of an overall interim assessment system.

IA Study Key Finding 2.3. In our sample, the most common method of administering the Smarter Balanced Interim Assessments was to the entire class.

Supporting Evidence:

- 87% of the test administrators who responded to our survey reported giving the assessment to the whole class.
- Less than 10% of the test administrators who responded to our survey reported giving the assessment to subsets of students in the class.

Only test administrators were asked to indicate how they administered the interim assessments. In our sample, 675 test administrators reported giving the assessments to all students in a class, 57 test administrators reported giving the assessments to the whole class and to subsets of students, and 50 test administrators assessed only students with disabilities. Very few test administrators reported giving the assessment to only remedial students, to only EL students, or to only advanced students.

Interim Assessment Study Research Question 3. To what degree were schools successfully prepared to administer the interim assessments (e.g., training materials were clear, system components worked)?

A multitude of resources are available to support the use of Smarter Balanced Interim Assessments. For example, LEA and school coordinators have access to (a) online Webinars, (b) guidance on how to use interim assessments to inform instruction, and (c) technical specifications for accessing the required systems. Some tools are required (e.g., TOMS, test administrator interface) and some would be used only if one accesses those features (e.g., Interim Assessment Reporting System and Hand Scoring System). We asked how helpful various resources and training opportunities were to stakeholders. Additionally, we asked LEA coordinators, site coordinators, and test administrators how easy it was to use TOMS, the Interim Assessment Viewing System, Test Administrator Interface, Online Reporting System (ORS), Interim Assessment Reporting System, Interim Assessment Hand Scoring System, and Administration and Registration Tools (ART).

LEAs in the state are still in the early implementation stages of the Smarter Balanced Interim Assessments component of the CAASPP System. In some ways, both CDE and Smarter Balanced are continuously making revisions and improvements to the available resources and system interfaces, as are LEAs continuously improving how they utilize the resources. These survey questions were designed to elicit general experiences; we did not explore in-depth what specific features of each were difficult or which features educators might have used or found more helpful. Even in instances where respondents perceived preparation and training to be difficult, there is no implication that efforts by the CDE and the Smarter Balanced Consortium are ineffective rather than in the process of development and improvement. Reported difficulties are relative to how users interpreted their experiences, regardless of whether they aligned with the intended use.

IA Study Key Finding 3.1. LEA coordinators generally found the various resources helpful and the various systems easy to use.

Supporting Evidence and Additional Information:

- Across the training resources listed as response options, approximately 50% of the LEA coordinators who responded to our survey found the resources to be helpful.
- The caaspp.org resource links were perceived as helpful by 63% of LEA coordinators who responded to our survey.
- Across the various Smarter Balanced Interim Assessment systems, 46% of LEA coordinators who responded to our survey found the systems to be easy to use.
- For our sample of LEA coordinators, the Test Administrator Interface was the easiest to use (73%).

LEA coordinators in our sample generally perceived the various preparation and training resources as helpful. No single resource emerged as the "most helpful." Although approximately 70 percent of LEA coordinators attended the Fall CAASPP Institute, only about half attended the Spring CAASPP Institute (54%). The two institutes were designed to be attended as a two-session training. While LEA coordinators generally reported that use of the various interim assessment-related systems was easy, they indicated the Interim Assessment Hand Scoring System was the most difficult to use (only 21% reported it was easy to use). School site coordinators generally reported that use of the various interim assessment-related systems was easy; the easiest system to use was the test administrator interface (84%).

IA Study Key Finding 3.2. Many test administrators in our sample perceived the CAASPP Interim Assessment User Guide to be helpful, although many other resources were reported as "not applicable."

Supporting Evidence:

- 64% of test administrators who responded to our survey reported the Interim Assessment User Guide to be helpful.
- Across a list of resources, an average of 69% of the test administrators who responded to our survey reported them as "not applicable," suggesting they did not attend CAASPP Institutes or webinars or did not consult resources such as CAASPP e-mail updates. [46]

[46] For Likert scale items presented in a matrix format with several rows of response options, we provide an average percentage or mean (M = x%) for the set of options.

Although the majority of test administrators did not use many of the available resources, the CAASPP Interim Assessment User Guide was used the most and was generally perceived as helpful. The majority of test administrators reported (M = 55%) not using various functions within the Interim Assessment Reporting System (e.g., viewing their students' results, downloading results for later analysis). Because LEA coordinators control the functions and permissions to the Interim Assessment Reporting System via the Administration and Registration Tools (ART) system, the high percentage of TAs that reported not using most of the functions may reflect lack of permission, rather than lack of use. Of those who tried to use the system, the percentage of those who found it easy to use was similar to the percentage that found it difficult.

IA Study Key Finding 3.3. Many LEA coordinators in our sample reported providing LEA-specific in-person training or LEA-specific online or print resources on the interim assessments.

Supporting Evidence:

- 60% of LEA coordinators who responded to our survey indicated they provided in-person training and 42% provide online or print resources.
- A majority of LEA coordinators who responded to our survey rated their LEA-specific resources and the caaspp.org resource links as helpful or very helpful in preparing for the Smarter Balanced Interim Assessments (60% and 63%, respectively).
- More test administrators who responded to our survey found the school-specific training/resources to be helpful or very helpful (45%), compared to those who found the LEA-specific training/resources to be helpful or very helpful (27%).

Based on our more detailed analysis, the size of the LEA might influence whether LEA-specific in-person training is offered. More coordinators in large LEAs reported offering in-person training (101 of 147 large LEAs, 69%) compared to coordinators in small LEAs (25 of 52 small LEAs, 48%).

Interim Assessment Study Research Question 4. To what degree is the information about test administration procedures, as included in interim assessment resources, followed?

Schools and districts have a great deal of flexibility in deciding whether, when, and how the interim assessments are administered. Information about how to select, administer, and monitor interim assessment test sessions; how to verify students' accessibility support settings; and how to understand the student testing interface is provided to LEAs in the *CAASPP Interim Assessment User Guide*.[47] We sought to evaluate the extent to which test administrators found the information useful and followed the procedures. In its training materials, the CDE describes the interim assessment component of the CAASPP System as occupying "a middle position between short-cycle formative assessment and long-cycle summative assessment."[48] Both non-standardized administration (assessment "for" instruction) and standardized/benchmark administration (assessment "of" learning)—recommended to enhance longitudinal analyses of results within LEAs or in support of using the tests as performance benchmarks—are appropriate and encouraged.

[47] The March 2016 version was prepared by the Smarter Balanced Assessment Consortium, with California customization by ETS.

[48] CAASPP Institute Module 3.1 Introduction to Interim Assessment, ~~http://www.cde.ca.gov/ta/tg/sa/introtointerim31.asp~~ [Note, the preceding Web address is no longer valid.]

IA Study Key Finding 4.1. Most test administrators in our sample reported that their school attempted to standardize how the Smarter Balanced Interim Assessments were administered (e.g., required formal training on some components and required that procedures outlined in the *CAASPP Interim Assessment User Guide* be generally followed).

Supporting Evidence:

- Over four fifths of the test administrators who responded to our survey (83% of 863) reported that they followed the *CAASPP Interim Assessment User Guide* procedures to a moderate or extreme degree; a much smaller percentage (16%) indicated teachers were allowed great flexibility in administration.

In comparison to test administrators, site coordinators reported wider experience with attempts to standardize how the assessments were administered. Responses varied from 32 of 86 coordinators stating that very formal procedures were required to 28 of the coordinators saying their site allowed great flexibility).

Interim Assessment Study Research Question 5. To what degree do LEAs perceive the interim assessments impact instructional practice and student achievement?

One of the main purposes of using the Smarter Balanced Interim Assessments is to obtain information on student progress toward mastery of skills, although other uses the CDE encourages are to allow educators to engage in professional learning, to deepen teacher content knowledge, and to observe students' use of the testing platform and accessibility supports. Because this was the first full year of Smarter Balanced Interim Assessment implementation, it is reasonable to expect some variety in the ways that LEAs use the interim assessments. As LEAs become more familiar with the available features, and as CDE and Smarter Balanced continue to make improvements, the ways in which LEAs perceive the interim assessments to be useful will likely change.

IA Study Key Finding 5.1. LEA and school coordinators and test administrators who responded to our survey reported that Smarter Balanced Interim Assessments were useful in familiarizing students and teachers with the format and content of the Smarter Balanced Summative Assessments, which are mandated by the state.

Supporting Evidence:

- The finding is consistent with analysis of statewide assessment data, which shows that schools that administered the Smarter Balanced Interim Assessments (either IAB or ICA) had more improvement in school-level scale scores than those that did not administer the interim assessments.
- Familiarizing students and teachers with the testing system, item types, and scoring rubrics was the most useful benefit reported by all three stakeholder groups (M_{LEA} = 72% of 285, M_{SITE} = 73% of 90, M_{TA} = 54% of 863).[49]

[49] For Likert scale items presented in a matrix format with several rows of response options, we provide an average (M = x%) for the set of options.

Smaller percentages of respondents report they are using results for improving instructional practice or student achievement. Although 49% of LEA coordinators report that *no* LEA-wide decisions about educator supports are made based on the Smarter Balanced Interim Assessments, there is some variety among those who do use interim assessments to make such decisions (99 coordinators report using interim assessments to inform department meeting topics, 70 coordinators report using them to inform professional development offerings). Using interim assessments to determine course placement or promote differentiated instruction were the least useful activities reported across all three stakeholder groups (M_{LEA} = 14%, M_{SITE} = 20%, M_{TA} = 23%).

The majority of LEA and school site coordinators did not question the validity of the Smarter Balanced Interim Assessment results. Test administrators did express somewhat less confidence that the results provided valid indicators of student achievement. Almost as many test administrators reported having little to no confidence (43%) as did those who reported having some to a great deal of confidence (56%) in the validity of the interim assessments.

Interim Assessment Study Research Question 6. What challenges existed in the 2015–16 school year that could be improved for 2016–17?

With 2015-16 being the first full year of implementation of the Smarter Balanced Interim Assessments, it is reasonable to expect some misinformation and higher percentages of challenges with the interim assessments in comparison to the summative assessments. Some of the challenges may stem from the newness of the variety of elements needed to interact with the assessments, as well as the related resources to support teaching and learning. LEA and site coordinators in our sample rated how challenging various functions and features of the Smarter Balanced Interim Assessment system were with respect to assessment administration, test content, hand scoring, and reporting. As the CDE and Smarter Balanced continue to make improvements and revisions, the ways in which educators and administrators perceive the various functions, components, and processes will likely change.

IA Study Key Finding 6.1. LEA and school site coordinators and test administrators in our sample did not report significant challenges with Smarter Balanced Interim Assessments.

Supporting Evidence:

- An average of 81% of LEA coordinators who responded to our survey rated various administration activities (e.g., bandwidth delays) as not challenging or only a minor challenge.
- An average of 61% of LEA coordinators who responded to our survey rated test content features (e.g., content was too difficult) as not challenging or only a minor challenge.
- An average of 46% of LEA coordinators rated reporting activities as not challenging or a minor challenge.

None of the stakeholder groups reported any major challenges with administration. They generally felt most issues were either a minor challenge or not a challenge at all. Most did not report issues with bandwidth, or difficulty setting up the assessment administration session. Major challenges reported by some LEA coordinators were determining the appropriate designated supports and accommodations (28%), scheduling time to access computers (19%), and unexpected system crashes (19%). School site coordinators noted as most challenging

student familiarity with technology (22%) and scheduling time to access computers (22%). The most challenging activities for test administrators were unexpected system crashes (23%), scheduling time to access computers (23%), and student familiarity with technology (20%).

School site coordinators and test administrators expressed general agreement that the Smarter Balanced Interim Assessments result in positive experiences for students with disabilities (SWDs) and English learners (ELs); however, they also expressed some level of uncertainty as to whether the designated supports and accommodations promote better access, familiarity, and comfort. Based on our more detailed analysis, more LEA coordinator survey respondents from LEAs with high EL populations (66% of 148) reported as a *major challenge* the adequacy of the supports and accommodations compared to LEA coordinator survey respondents from LEAs with low EL populations (44% of 94).

None of the stakeholder groups reported test content being too easy as a major challenge. Some LEA coordinators (21%) reported the opposite—that the content is too difficult— and more school site coordinators and test administrators reported difficulty was a major challenge (46% and 47%, respectively).

Although the hand scoring activities of Smarter Balanced Interim Assessments can facilitate professional learning for educators and help them better understand the system of scoring—including awareness of rubrics that can be used in the classroom for teaching and student learning— some respondents in our sample cited the hand scoring aspect of interim assessments as a challenge. During LEARN meetings, focus groups, and interviews, the topic of hand scoring was described by some LEA and school staff as an obstacle with the Smarter Balanced Interim Assessments. The effort to organize scorers, understand and implement the training procedures, and conduct hand scoring, as well as the costs associated with hand scoring, were perceived as insurmountable by some LEAs and schools. These concerns, which were identified during our qualitative data collection phase, are supported by our survey results. A large number of LEA and school site coordinators reported that (a) committing adequate time to complete hand scoring (LEA: 48%; SITE: 35%) and (b) identifying classroom teachers who are willing to hand score their own students' responses (LEA: 33%; SITE: 32%) were the greatest challenges. Based on the relatively high percentages of LEA and school site coordinators who indicated they did not know whether certain activities were a challenge or not (M_{LEA} = 39%; M_{SITE} = 42%), it is possible many CAASPP coordinators have not yet been exposed to some of the pertinent hand scoring components (e.g., navigating the Hand Scoring System). Similarly, a large number of test administrators reported that they did not know whether various hand scoring activities were challenging (M_{TA} = 60.2%), suggesting that they might not have had any or much experience with hand scoring.

Across all three stakeholder groups, challenges with reporting results were consistently indicated. LEA coordinators reported that the adequacy of detail in results (not reported by target, CCSS, or strand) to inform changes to instruction was the greatest challenge (57%). Based on our more detailed analysis, this concern was cited more frequently as a *major challenge* for larger (68% of 137) than smaller (35% of 48) LEAs in our sample. These findings reflect user perceptions absent an understanding of the psychometric limitations of reporting results based on very few test items. Other significant challenges included lack of integration with other student record systems (37%) and difficulty aggregating/grouping student scores (32%). Although school site coordinators likewise indicated the inadequacy of detail in results (not reported by target, CCSS, or strand) to inform changes to instruction as a major challenge (36%), many seemed to lack experience with reporting activities (as indicated by large percentages who stated they did not know whether challenges existed: 41%). For the most part, test administrators indicated they did not know whether challenges existed (M_{TA} = 59%),

suggesting they either did not have much experience with these activities or have no opinions on those components.

Based on the key findings from the Smarter Balance Interim Assessment Administration Study, we offer the following recommendation:

IA Study Recommendation: Continue to monitor the various Smarter Balanced Interim Assessment systems and components. We recommend that the CDE also continue collecting feedback from schools and LEAs, as well as from other Smarter Balanced Assessment Consortium members, to see where reasonable improvements could be made to the system.

ETS collects and reports information on use of interim assessments in TOMS. It may be possible to modify or expand that system slightly to allow monitoring of additional aspects of interim assessment usage.

CDE and Smarter Balanced provide to administrators and educators a wide variety (e.g., videos, manuals, e-mail updates, online links, responses to Frequently Asked Questions) and quantity of resources and tools that continue to develop and improve. Because this is the first full year of implementation of the Smarter Balanced Interim Assessments, and because not all LEAs have fully mastered the complexity of the overall CAASPP System or fully implemented all of its components, it is not surprising that LEAs are still learning how to use and take advantage of the capabilities of the interim assessments. As LEAs become more familiar with the available features and as CDE and Smarter Balanced continue to make improvements, the ways in which LEAs perceive the interim assessments to be useful will likely change. Additionally, a shared responsibility of understanding and implementation among CDE, Smarter Balanced, and LEAs exists in ensuring successful implementation of the Smarter Balanced Interim Assessments. Therefore, HumRRO recommends continuing to monitor these efforts to ensure that changes occur in the predicted direction. We particularly recommend continuing to monitor the following:

a) *How the different types of Smarter Balanced Interim Assessments are used, particularly to inform instruction*: ETS can provide aggregated counts of interim assessment administrations, but little information is systematically collected about how using interim assessments impacts instruction. Our study indicated that many educators were using the interim assessments as a way to familiarize students and teachers with the Smarter Balanced Test Administration Interface and item types (IA Study Key Finding 1.1). This is reasonable to expect in the early implementation stages, and with millions of interim assessments administered in 2015–16, there is evidence that educators are using the interim assessments in large numbers. As LEAs transition to fuller implementation, however, we would expect to see the interim assessments being used more closely with their original intent of gauging student progress toward meeting content mastery. Examining when during the school year teachers administer the interim assessments and the manner in which they are administered (current test settings capture non-standardized versus standardized administration) can provide useful information about how they might be impacting instruction.
b) *The use of the various training and preparation resources and tools, particularly among test administrators*: Findings from all three surveys suggest that many school-level and LEA-level educators in the HumRRO sample are not using many of the

Smarter Balanced Interim Assessment resources and systems (IA Study Key Finding 3.2). This could be an awareness issue in that LEAs and teachers just have not yet had the time to learn about all of the new features.

c) *The satisfaction with the content and usefulness of interim assessment results*: Many of our respondents indicated a desire to have more details in the interim assessment score reports, for example item or target level results (IA Study Key Finding 6.1). Although this might be a desirable feature, it is contrary to expert psychometric recommendations of reporting at the subscale level with too few test items. We did not investigate whether respondents would be willing to tolerate a longer test in order to support more detailed reporting. A dissatisfaction with the available reporting options manifested throughout the survey findings, impacting respondents' perceptions of how easy the Smarter Balanced systems were to use, perceived challenges, and desired improvements. Monitoring how these perceptions evolve (or not) will help indicate whether educators are becoming more familiar with how the Smarter Balanced Interim Assessments are intended to be used and more pleased with the available level of score reporting.

Results from Access to Designated Supports and Accommodations Study (Chapter 4)

The specific purposes of this study were to (a) examine the availability and use of testing supports and accommodations for students with disabilities (SWDs) and English learners (ELs) on the Smarter Balanced ELA and mathematics assessments and (b) determine whether the tools used for these assessments are consistent with those used routinely by students in their classrooms. Our study obtained a sample of LEAs, schools, and local staff that was about one-half of the originally targeted sample size (6 of 12 LEAs and 15 of 36 schools). The reduced sample sizes allowed us to collect information from the study participants in greater depth, leading to what was more like a series of case studies. Although there are significant limitations with respect to the generalizability of this study's findings, results from the study may inform future larger-scale studies on this topic.

We generated qualitative and quantitative data to address the study's research questions. The qualitative data were generated from phone-based focus groups, in-person interviews with local staff during school site visits, and observations of instruction and assessment. Quantitative data related to the use of instructional accommodations were generated from teacher reflection checklists and from items on the survey conducted as part of the Interim Assessment Administration Study (see Chapter 3). A full description of the development of data collection instruments, training of interviewers, study participants, and other study details and analyses can be found in Chapter 4.

Although the study sample includes a range of LEA and school characteristics, we cannot be sure if there were unmeasured characteristics that are related to whether or not an LEA opted to participate in this study. For example, there was anecdotal indication that some LEAs were hesitant to participate if they felt their staff lacked familiarity with issues related to testing accommodations. Participation within LEAs also proved to be a challenge. For example, only about half of teachers who agreed to complete reflection checklists actually did so. The following key findings based on these limited samples represent the LEAs, schools, and teachers who participated in the study, but may not generalize precisely to the full state population.

Study results reflect the combined data from focus group interviews, site visit observations and interviews, teacher reflection checklists, statewide interim assessment survey data, and 2015–

16 summative assessment data on accommodations and designated supports. Below we summarize for each research question our key findings and the evidence that supports them.

Access to Designated Supports and Accommodations Study Research Question 1. Is the general assessment accessible to moderately disabled students and English learners through the provision of accommodations and supports?

The Smarter Balanced Summative Assessments include a wide variety of universal tools, designated supports, and accommodations designed to allow students to access the assessments and demonstrate what they know and can do. Within the new online testing environment, the entire suite of sophisticated designated supports and accommodations that students can be offered is a tremendous expansion beyond what has been available in paper and pencil mode.

The online CAASPP Smarter Balanced assessments allow universal tools, designated supports, and accommodations (embedded and non-embedded). ***Universal tools*** are available for all students based on student preference and selection, and include, among other resources, breaks, digital notepad, English glossary, highlighter, strikethrough, zoom, and mark-for-review. ***Designated supports*** are available to *all* students when determined for use by an educator or group of educators (with parents/guardian and student input, as appropriate) or specified in the student's individualized education program (IEP) or Section 504 plan. These include but are not limited to color contrast, masking, text-to-speech, ability to turn off any universal tool, separate setting, special lighting or acoustics, and administration of the test to the pupil at the most beneficial time of day. ***Accommodations*** are available if specified in the student's IEP or Section 504 plan and include braille; text-to-speech for ELA reading passages alternate response options; American Sign Language for Writing, Listening, and Mathematics; print on demand; read aloud for ELA reading passages; scribe for Writing; and speech-to-text.[50]

Access Study Key Finding 1.1. Students with disabilities and English learners were offered a wide range of accessibility features during interim and summative assessments, although actual use of these features cannot be determined.

Supporting Evidence:

- Population-level data on summative assessment accommodations indicate that all students with disabilities and English learners were offered at least one accessibility feature through the provision of accommodations, designated supports, and universal tools. The assessment software does not capture student use of features, however.

Population level data from the 2015–16 summative assessment administration indicate that all students with disabilities and English learners were offered at least one accessibility feature through the provision of universal tools, designated supports, or accommodations (see Table 4.8). The large majority had access to embedded universal tools, unless access was turned off as part of a designated support.[51] The percentages of students receiving math designated supports is high for students with disabilities, English learners, English learners with disabilities,

[50] Source: http://www.cde.ca.gov/ta/tg/ai/caasppmatrix1.asp (Retrieved 10/5/16).
[51] By default, universal tools are available to all students. If needed, universal tools may be turned off. Turning off universal tools is a designated support.

and students not identified in either of these subgroups due to the fact that nearly all students were offered the translations (glossary) designated support (available for math only).However, data on the actual use of these features by students are not available.

The statewide interim assessment survey asked questions directly related to the accessibility of the Smarter Balanced Interim Assessments. Though the survey findings are not generalizable statewide, the variation in survey responses in our sample suggest that site coordinators and test administrators had mixed views on whether accessibility features increased student access to the interim assessments (See Figures 4.1 and 4.2). School site coordinators generally agreed the accommodations and designated supports on the interim assessments contributed to positive experiences for SWDs. Their level of agreement to the statements that the interim assessment accommodations and designated supports improved access to the test (66% indicated they strongly agreed or agreed) and that the accommodations and designated supports functioned properly during testing (64% indicated "strongly agree" or "agree"). In contrast, less than half of test administrators agreed that for SWDs, the accommodations and designated supports (a) improved access to the test, (b) are consistent with instructional practices used in the classroom, and (c) functioned properly during operational testing. Additionally, a large percentage of test administrators reported not knowing whether the Smarter Balanced Interim Assessments resulted in positive experiences for SWDs (as indicated by responding "I don't know"). Test administrators also indicated uncertainty regarding whether the accessibility features of the interim assessments contributed to positive experiences for ELs.

Data from focus groups and school visits, though limited to responses from staff and teachers in six relatively diverse LEAs, suggest that teachers believe most students are able to demonstrate what they know and can do through the provision of designated supports and accommodations; however, responses varied across schools. The primary concern teachers and LEA staff expressed regarding accessibility is with the integration of computers in schools. LEA and school staff were asked if they felt the summative and interim assessment scores are valid and reliable and the responses were mixed, even within the same LEA. Four LEAs reported (5 separate comments) that they felt the assessment scores do reflect what their students can do. However, we received 10 comments from LEA staff or teachers from 4 LEAs indicating that assessments may not allow accurate inferences regarding what students know and can do.

Access Study Key Finding 1.2. In these early days of the new assessment environment, some educators in our sample expressed a lack of confidence or knowledge about CAASPP accessibility procedures, although resources for learning about these procedures are available.

Supporting Evidence:

- Most teachers from six LEAs who participated in focus groups and interviews reported that they had received minimal, if any, training on the accessibility features of the Smarter Balanced assessment system.
- Some teachers from six LEAs who participated in focus groups and interviews reported that students' familiarity with Smarter Balanced accessibility features may be limited due to lack of computer resources and/or practice time.

During our focus groups and interviews, many of the issues raised by participants were related to familiarity with and understanding of CAASPP accessibility policies, procedures, and tools. Several educators indicated they were not familiar with all allowable accommodations or

expressed concerns they would unknowingly administer an accommodation or support in a manner that invalidated student scores. Some staff responsible for assigning designated supports and accommodations within TOMS found the process cumbersome and may have lacked full understanding of the most efficient ways to implement processes and procedures. Several teachers of students who communicate via American Sign Language (ASL) expressed concerns about how to appropriately address student questions when they were unfamiliar with or unsure about the signing presented in ASL videos. Informational resources are available through both the CAASPP and Smarter Balanced Web sites and we anticipate that district- and school-level staff will increase their awareness and understanding as the CAASPP system continues to be implemented.

The most common concern impacting student scores (across grade and disability) is the lack of familiarity with the computer-based testing environment. This concern was raised in all six LEAs either during observation, school interviews, or focus groups regarding participants' familiarity with Smarter Balanced universal tools, designated supports, and accommodations. This observation applies across the board, to CAASPP coordinators, teachers, test administrators, and students. For example, classroom teachers, who typically served as test administrators, frequently indicated that they had received minimal, if any, training on the accessibility features of the Smarter Balanced assessment. Similarly, focus group participants indicated that student familiarity with the accessibility features may have been constrained due to limited availability of computer resources and classroom instructional time.

In addition to computer related issues, educators in three LEAs that participated in focus group and interviews raised concerns that the rigor of the Smarter Balanced test items was set too high for SWDs and ELs in terms of required reading comprehension level and overall item difficulty. Also, in three of the LEAs, through assessment observation and staff interviews, it was noted that some students had difficulty knowing how to answer the various types of Smarter Balanced items, primarily in math (e.g., drag and drop, hot spot, table fill in, graphing, equations).

Access to Designated Supports and Accommodations Study Research Question 2. To what extent do the supports and accommodations provided and used in the interim and summative assessments match those used in classroom instruction for individual students?

The *Smarter Balanced Resources and Practices Comparison Crosswalk* demonstrates that the universal tools, designated supports, and accommodations have been aligned to a wide variety of practices that teachers report using on a regular basis during instruction and classroom assessment. Using a checklist developed by HumRRO and based on the *Smarter Balanced Resources and Practices Comparison Crosswalk*, teachers in our study sample documented their instructional practices and classroom assessment (i.e., formative, diagnostic, interim, or benchmark testing) for individual students. They also recorded accommodations they used with these students over the course of the 2015–16 school year. HumRRO then merged teacher checklist data with data on summative assessment accommodations and designated supports and created indicators of when students received both an instructional practice and the corresponding summative assessment accessibility feature. These data are supplemented with focus group and site visit data, including eight observations of summative assessment administration.

Access Study Key Finding 2.1. In our sample, assessment accessibility features offered to students were generally aligned with instructional practices of the students' teachers.

Supporting Evidence:

- Among the most frequently used instructional practices indicated in the small matched sample of students, the majority of students typically were offered the associated accessibility feature during assessment, most often through the provision of universal tools.

Teachers who completed reflection checklists reported using many of these practices during instruction. Completed checklists were received from 15 teachers who reported the use of instructional and testing accommodations for a total of 42 students. Among the students for whom data were reported, 16 (38%) were classified as SWD, 10 (24%) as EL and 12 (29%) as both SWD and EL. No classification information was provided for 4 (10%) students in the sample.

Among the small sample of students for whom data on both instructional practices and summative accommodations/designated supports were available, the majority (28 of 34 for ELA; 32 of 34 for math) were offered summative accommodations, designated supports, or universal tools that corresponded with at least half of the instructional practices reported by teachers (See Table 4.9). The most commonly used instructional practices reported by teachers were available to students during the summative assessment through the provision of universal tools. Some teachers reported using an instructional practice throughout the year for some students who were not offered the corresponding summative accommodation or designated support during the assessments. For example, teachers reported using multiplication tables with several students, but those students were not offered the corresponding multiplication table non-embedded designated support on the math test.

Focus group notes and data from site visits indicate that CAASPP coordinators and resource teachers worked to ensure the accommodations offered to students during summative assessments reflected the instructional practices and supports to which students were accustomed. Site visitors in one LEA observed teachers and CAASPP coordinators working together to review testing accommodations settings prior to starting the test to ensure the intended accommodations were offered, but this practice did not appear to be widespread.

Access to Designated Supports and Accommodations Research Study Question 3. Are there types of supports or accommodations used by students when learning in the classroom that are not used on assessments?

Access Study Key Finding 3.1. In our sample, teachers rarely reported using instructional practices that were not reflected in the summative assessment accommodations, designated supports, and universal tools available.

Supporting Evidence:

- When given the opportunity to identify "Other" instructional practices, teachers in our sample who completed checklists tended to identify practices that were already listed, or that could not be provided during testing without changing the construct.

Access Study Key Finding 3.2. In our sample, several educators identified issues with the braille accommodation that may impact the correspondence between students' instructional and testing experiences.

Supporting Evidence:

- During site visits, educators from one LEA noted difficulties in efficiently administering the braille accommodation in tandem with the computer adaptive testing format of the ELA assessment.
- During site visits, educators from one LEA noted that assistive devices and other equipment that students routinely use in the classroom were not currently compatible with the Smarter Balanced testing platform.

Two specific issues were noted involving students tested with a braille accommodation. First was the amount of time and resources needed to produce test items in braille format when test content is delivered using the online test delivery system with access to the Computer Adaptive Testing (CAT) algorithm. Observers noted students were required to wait several minutes in mid-test for items identified via the algorithm to be printed in braille format. Administering the braille accommodation via fixed test form, which is available for mathematics, would be a simple way to resolve this issue, but some teachers voiced concern that the fixed form testing experience was dissimilar from that of general education students. The second issue involved differences in technology such as refreshable braille displays and electronic braille note-takers that students use on a regular basis versus those that are compatible with the Smarter Balanced testing platform. These issues could potentially impact student motivation and attention, and ultimately their performance.

The *Smarter Balanced Resources and Practices Comparison Crosswalk* introduced in the discussion of Research Question 2 (see Table 4.5), presents a summary of the usage of instructional practices and classroom testing supports. The table shows that many practices are used during instruction only for a substantial percentage of students. This may be indicative of supports and accommodations that are effectively used during full class or one-on-one instruction to introduce new concepts or review prior knowledge but that are less appropriate or not allowed when students are asked to demonstrate their knowledge independently during classroom assessments.

Teachers completing the reflection checklists were also given the opportunity to list instructional practices that they commonly use but which were not part of the *Smarter Balanced Resources and Practices Comparison Crosswalk*. Eight teachers reported additional instructional practices for a total of 25 students. The most commonly listed of these instructional practices were graphic organizers, which are not listed in the *Crosswalk* and thus do not have a corresponding summative assessment accessibility feature. However, most of the "other" instructional practices listed by teachers (e.g., interactive whiteboard, smaller setting) do correspond with a summative assessment accessibility feature or otherwise would be classified as an unlisted resource that could potentially impact the construct being measured.

There was evidence from the focus groups and school visits that teachers and CAASPP coordinators (a) were not familiar with all of the available testing designated supports and accommodations and (b) were not fully knowledgeable about which accommodations and designated supports were appropriate for their students. For example, when asked about

stacked translations for EL students, focus group participants from 3 LEAs that included EL students indicated they were not at all familiar with this designated support.

Also, data from focus groups and site visits suggest that moderately disabled students and ELs do receive supports during instruction that are not provided during testing. In several cases, this discrepancy may be appropriate because the supports would likely modify the construct being measured by the test. Teachers indicated these supports are not only important for instructional delivery but they are important to help teachers understand what students know and can do. These supports include, but are not limited to: coaching, referencing previous content, graphic organizers, sentence frames, modeling, questioning, repetition, and visual checklists. When asked to demonstrate their knowledge or skills, students often require some prompting, reminders of previous lessons, or confirmation they are on the right path in order to persevere on a task. For example, one student with a disability was unable to remember how to answer the first math question on the summative assessment. He asked for assistance and was told to do his best. He spent the entire testing time clicking repeatedly on every option he could find to either get help solving the problem or to skip it. In the interview associated with this observation, the teacher expressed frustrations that this student's test score will be zero even though math is one of his strengths. Another teacher stated that EL students might have difficulty understanding the meaning of common words when they are used figuratively, so they need some clarification during assessments. However, there is an embedded glossary available, in up to 10 languages, to meet this need. Again, this level of support may not be appropriate within a testing context, depending on the construct being measured.

Access to Designated Supports and Accommodations Research Study Question 4. How often do students attempt to use test supports and accommodations that they do not use in classroom instruction?

Research Question 3 addresses gaps in the correspondence between summative testing accommodations and the instructional supports students regularly use to access content and demonstrate learning in the classroom. Research Question 4 addresses issues related to students accessing allowable summative assessment accommodations with which they are not familiar.

Access Study Key Finding 4.1. There does not appear to be a widespread issue of students being offered large numbers of unneeded accommodations and designated supports on the summative assessments.

Supporting Evidence:

- Population-level data from ETS shows that students with disabilities and English learners were typically offered 1–2 accommodations or designated supports. Very few students were given large numbers of accommodations or supports.
- The sample of educators and testing coordinators who participated in our focus groups tended to express wariness of offering accommodations that could be considered inappropriate and thereby might invalidate students' scores.

Access Study Key Finding 4.2. Teachers and district staff from all six LEAs in our sample expressed concerns about the impact of limited computer experience on students' use of computer-based accessibility features.

Supporting Evidence:

- Teachers participating in focus groups, interviews, and site visits frequently cited students' (particularly younger students') lack of experience with computers in general and with the Smarter Balanced platform in particular, as potentially impacting student test performance.

A common theme among district and school staff was that students lacked computer and keyboarding skills considered essential to performance on the summative assessment. Several school and district staff expressed concerns about students' lack of experience with certain computer-based tools, and with the Smarter Balanced platform, and how this might impact test performance.

As noted in Table 4.10, there are few classroom assessment accommodations that are not also used during instruction, which provides some evidence related to the alignment of instructional and classroom assessment practices. For the majority of practices listed on the reflection checklist, no teachers indicated they are used only for classroom testing. The practice with the highest percentage of students receiving it only during testing (14.3%) was *simplified test directions*. Other practices or supports used only during testing were reported for fewer than 10 percent of students, and the majority of practices and supports were not indicated as used during testing only for any students.

HumRRO school visitors found no indication that students were intentionally being given access to designated supports or accommodations that were not appropriate for them or different from their typical classroom instruction and IEP-designated needs. Rather, it was more likely that LEA and school staff responsible for assigning testing accommodations would err on the side of caution. For example, some school CAASPP coordinators and teachers expressed concern over the implications of providing access to a designated support or accommodation that was not appropriate, or of unintentionally providing a non-embedded support inappropriately, and thereby invalidating a students' test score.

Although HumRRO site visitors did not report any serious, widely experienced problems among students who were observed using designated supports or accommodations during summative assessment, there were indications that some students in 3 schools (2 LEAs) were not completely comfortable with the Smarter Balanced platform. For example, some students tested in mathematics were observed having difficulty understanding the instructions for how to type answers using available math notations and how to use math tools such as rulers and protractors. Also, one teacher noted that through classroom instruction students have come to understand the benefits of tracking meaningful text elements through highlighting and underlining, but a lack of familiarity with computer-based highlighting and underlining tools hindered the student's use of these tools during interim and summative testing.

Differences in presentation or delivery mode of supports during assessment were of particular concern among teachers of students with vision and hearing impairments. Teachers and staff in 3 LEAs noted that some visually impaired students had difficulty with computerized voices or found difficulty using different braille reading hardware/software that was not compatible with the Smarter Balanced platform. Further, teachers and HumRRO observers noted that the process of printing test

items in braille format is very time consuming and requires students to wait several minutes in mid-test before they can access the next test item. Administering the braille accommodation by fixed form (i.e., paper format), which is available for mathematics, would be a simple way to resolve this issue, but some teachers voiced concern that the fixed from testing experience was dissimilar from that of general education students. Teachers also noted that students experienced difficulties using a braille reader during practice testing because passages displayed only one line at a time making it difficult to reread and review long passages. Similarly, some students experienced text underlining that did not appropriately render. Additionally, some students needed different color overlays or color contrast for particular test items and teachers indicated it was cumbersome within the Smarter Balanced testing platform to make these changes at the item level. A major concern among teachers of deaf students was the use of signed English rather than ASL on the summative assessment. Teachers noted the differences between ASL and signed English, and the lack of familiarity with particular signs used in videos on the summative assessment could mask conceptual understandings that students have achieved.

Teachers and staff in all 6 LEAs indicated that both teachers and students need to find more time to practice and use the Smarter Balanced system, or similar computer-based systems. Doing so, however, will take time out of current instructional time and will further draw on school or district resources to provide adequate numbers of computers. One LEA focus group summarized the concerns by stating that, in terms of summative assessment materials, there is not much disconnect; however, the issue seems to be with teacher and student understanding of the available tools, designated supports, and accommodations, and how they are used. Given these findings, the desire to use practice tests and training tests to familiarize students with the testing environment, and to reinforce with interim assessments as noted in our other study this year (Chapter 3), seems entirely appropriate, particularly for SWDs and ELs.

Based on the key findings from the Access to Designated Supports and Accommodations Study, we offer the following recommendation:

Access Study Recommendation: Continue to monitor the accessibility features available in the Smarter Balanced assessment systems to see where reasonable improvements could be made. Continue to collect feedback from stakeholders to identify and promote best practices for implementing accessibility features.

CDE and Smarter Balanced provide a wide variety (e.g., videos, manuals, the *Individual Student Assessment Accessibility Profile (ISAAP tool),* the accessibility matrix) of resources for district staff and teachers to use to support all students' access to assessments; however, teachers are still learning how to appropriately take advantage of the available universal tools, designated supports, and accommodations. Our study findings are generally consistent with what would be expected in the context of a newly implemented assessment system. As training and information dissemination increases, teachers' confidence and knowledge about CAASPP accessibility procedures is expected to increase. HumRRO recommends continued monitoring to ensure that changes occur in the predicted direction. We particularly recommend continuing to monitor the following:

a) *Trends and patterns in the universal tools, designated supports, and accommodations offered to students*: Though we found some evidence that students with disabilities and English learners are being offered access through these features, the assessment data collected from ETS do not include whether the features are *used* by the students (Access Study Key Finding 1.1). We encourage the CDE to work with Smarter Balanced to collect data and report on actual use of the features.

b) *How the training and other resources for accessibility features are being used:* There would be great value in learning from LEAs that have successfully implemented this aspect of the CAASPP System, in particular for students who are visually impaired and students who are deaf. Focus groups and interview findings from our small sample suggest that school-level and LEA-level educators have not yet achieved mastery of appropriate accessibility settings for Smarter Balanced assessments (Access Study Key Finding 1.2). This could be an awareness issue in that some LEAs and teachers have not yet had the time to learn all about accessibility features, or it could indicate a need for improvement in the training resources. It may be that teachers have a general awareness but haven't yet determined whether, for individual students, particular accessibility features are useful. Consider supplementing available professional learning resources by identifying and promoting best practices, for example conducting a post-test evaluation using a tool such as the After-test Accessibility Questions from the updated CCSSO Accessibility Manual (2016).[52]

c) *Use of the available Smarter Balanced assessment components:* Practice tests and interim assessments are known to be helpful for experimenting with various accessibility supports and in familiarizing students with disabilities and English learners with the online testing environment (as noted in IA Study Key Finding 1.1).

Results from Observations of CAASPP Training Sessions (Chapter 2)

Based on our observations, we summarize key findings and make one recommendation for professional development resources for the CAASPP System:

Training Sessions Key Finding 1. The CAASPP Institutes and Post-test Workshop provide high quality in-person training and valuable reference materials for remote access by teachers and other educators.

Supporting Evidence:

- Presenters were knowledgeable and energetic in communicating substantive content about targeted aspects of the CAASPP System to a variety of types of educators (e.g., assessment directors, principals, ELA or mathematics teachers, and content specialists) with differing perspectives on the CAASPP System and various levels of training and experience in assessment theory and implementation.
- Materials included a variety of graphics and other content organizers to illustrate and emphasize essential concepts.
- The CDE posted links to electronic files of all materials from these sessions.

HumRRO found these training sessions for LEA and school level staff to be very well designed and excellently managed and conducted. Agendas maintained participant engagement, provided opportunities for new ideas to be absorbed, and forged connections between staff from different LEAs by interspersing small group discussions and large group question-and-answer periods. In addition to delivering guidance about Smarter Balanced components, CAASPP Institutes addressed head-on the types of challenges educators experienced when

[52] CCSSO Accessibility Manual: How to Select, Administer, and Evaluate Use of Accessibility Supports for Instruction and Assessment of all Students

implementing the CAASPP System and provided opportunities for LEA and school staff to share their successful strategies. High quality supporting materials were distributed to participants and proved to be valuable reference documents that participants could take back to share with local staff who did not attend the training. The CAASPP Institute Train-the-Trainer binder, given to participants in electronic and hard copy format, included not only the presentation slides but also annotations for delivering the presentation; this was a particularly powerful resource with the potential for widespread effectiveness. The Post Test Workshop was a comprehensive session for leading teachers and district staff through a logical sequence of steps from data acquisition through a "research, recall, reflect and respond" process, with the goal of helping LEAs customize their own approach to using summative assessment results in the instructional cycle. The planning tools, including a data analysis process template with guiding questions and a sample completed template, were very well designed.

Training Sessions Key Finding 2. Participants attending training sessions arrive with a variety of levels of starting knowledge about the CAASPP System.

Supporting Evidence:

- Participants were from LEAs and schools that differed in their readiness for implementation of one or more components of the CAASPP System (e.g., were still in the "awareness" stage or early "transition" stage or were approaching full implementation of all components).

Though the facilitators of the CAASPP Institutes used the available time for training efficiently and led activities to support the predefined learning goals, some participants who had little prior training or came from schools or LEAs that differed in their readiness for implementation of one or more components of the CAASPP System appeared somewhat overwhelmed with the volume of content to digest. For participants from schools or LEAs that were still in the "awareness" stage (e.g., little or no use of the Smarter Balanced Interim Assessments) or "transition" stage (e.g., only sometimes used accessibility supports during summative assessments for all students, including SWD and ELs) , meeting the success criteria of the session (e.g., "Explain the connection between the summative assessment claims, the assessment targets, and California's new standard; and read and understand the value and use of the test blueprints") would require additional time reviewing and digesting the available training modules. Over time, as the CAASPP Institute training materials and other Smarter Balanced assessment resources, Webcasts, and manuals become more fully utilized, the knowledge gaps among educators across the state regarding such topics as how to provide access to supports and accommodations during testing and how to use and hand score interim assessments would be expected to diminish.

Training Sessions Key Finding 3. Operational training for hand scoring the Smarter Balanced Summative Assessments provides valuable professional development for teachers and curriculum specialists.

Supporting Evidence:

- Several educators attending the training for professional development purposes stated that the training was of great value and lessons learned would be shared with fellow LEA or school staff to help students improve their open-ended responses on future summative assessments.

HumRRO found the hand scoring training sessions very well managed, professionally conducted, and appropriately monitored to ensure the security of confidential materials. Discussions were collegial and thorough, and HumRRO observed facilitators using several techniques that were effective in ensuring appropriate application of the scoring rubrics. For example, participants were asked to explicitly describe how a response under discussion matched the scoring guidance. Although the specific purpose of these sessions was to certify raters seeking to be hired by one of the two testing contractors, MI or ETS, HumRRO learned that some public school educators who were not interested in rating for pay participated in the training because they were interested in learning more about the concepts and processes for scoring open-ended student responses on the summative assessments. In discussions with several of these educators attending for professional development purposes, we heard that the training was of great value and what was learned would be shared with fellow LEA or school staff to help students improve their open-ended responses on future summative assessments. For example, an educator from an LEA with a large population of English learners (ELs) was surprised and pleased to learn that California students who supported their mathematics responses with narrative in Spanish would have the responses scored by a certified Spanish Scoring Leader. The educator indicated that next year, newcomers with strong mathematics and Spanish skills but weaker English language skills would be directed to give their text response in Spanish along with the mathematics to have a better chance of getting credit for their explanations. Another educator from an LEA with a large population of ELs, on the other hand, was disappointed to learn that the same ELA scoring rubrics were used for all students, including ELs and students with disabilities (SWDs).

Training Sessions Recommendation: Continue to support professional development opportunities and maintain online resources that enhance LEA and school staff understanding of how best to utilize all components of the CAASPP System to improve teaching and learning.

Continuing to offer high quality training sessions, such as the CAASPP Institutes and CAASPP Post-test Workshops, will promote consistent, ongoing implementation of the CAASPP System and achievement of the system's goals in schools and LEAs across the state. Maintaining links to the great variety of training resources and materials on the CDE's Web site and on the caaspp.org site will provide teachers, administrators, and LEA staff throughout the state with easy access to information about the fundamentals as well as the intricacies of the complex CAASPP System. One suggestion for improvement that may assist educators from LEAs still in the transition stage of implementing CAASPP components (e.g., not yet administering interim assessments or not yet hand scoring any, not yet making full use of designated supports) is to create a roadmap that organizes available resources into a simplified set of starting points to develop foundational knowledge, which can later be expanded into more in-depth applied knowledge.

General Summary

Our study of interim assessments found wide usage of the CAASPP interim assessments. Initial use of these assessments focuses on familiarizing students with the summative test and testing environment. The use of interim assessment results to improve instruction will likely increase over time.

Our study of supports and accommodations found good alignment of supports and accommodations during testing for SWDs and ELs with those used routinely in classroom

instruction. A chief concern was that some students may not understand how to respond to some of the newer item types. Given this concern, the use of the interim (and practice) tests to familiarize students with the testing environment is entirely appropriate.

Overall, results to date indicate the California public school system has implemented the Smarter Balanced assessments in a way that provides stable and meaningful results. Current evaluation work has resulted in a number of suggestions for improving the efficiency and efficacy of the use of interim assessments as well as of the supports and accommodations offered to students during the assessments.

REFERENCES

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

California Department of Education. (2016). *Smarter Balanced Interim Assessments Web site.* Retrieved June 15th, 2016, from http://www.cde.ca.gov/ta/tg/sa/sbacinterimassess.asp.

Crawford, L. (2007). State testing accommodations: A look at their value and validity. New York: National Center for Learning Disabilities. Available online at http://www.ld.org.

Elliott, J., Thurlow, M., Ysseldyke, J., & Erickson, R. (1997). Providing assessment accommodations for students with disabilities in state and district assessments (Policy Directions No. 7). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available online at: http://education.umn.edu/NCEO/OnlinePubs/Policy7.html.

Hardoin, M.M., Becker, D.E., and Wise, L.L., eds. (2015). *California Assessment of Student Performance and Progress (CAASPP): 2015 Independent Evaluation Report* (2015 No. 044). Alexandria, VA. Human Resources Research Organization.

Olson, J.F. (2010). Guidebook for Studies of the Validity of Test Results for Test Accommodations. Washington, DC: Council of Chief State School Officers.

Shyyan, V., Thurlow, M., Christensen, L., Lazarus, S., Paul, J., and Touchette, B. (2016). *CCSSO accessibility manual: How to select, administer, and evaluate use of accessibility supports for instruction and assessment of all students.* Washington, DC: CCSSO.

Sireci, S.G. (2004). Validity issues in accommodating NAEP reading tests (Research Report 515). Amherst, MA: University of Massachusetts Amherst. Available online at: ~~https://www.nagb.org/publications/reports-papers/special-populations/sd-lep-conference-papers.html~~ [Note, the preceding Web address is no longer valid.]

Sireci, S.G., Li, S., & Scarpati, S. (2003). The effects of test accommodation on test performance: A review of the literature. Center for Educational Assessment (Research Report No. 485). Amherst, MA: University of Massachusetts at Amherst.

Thurlow, M.L. (April, 2007). Research Impact on State Accommodation Policies for Students with Disabilities. NCEO, University of Minnesota.

U.S. Department of Education, National Center for Education Statistics. (2003). NAEP Validity Studies: The Validity of Oral Accommodation in Testing, NCES 2003–06, by Timothy J. Weston. Project Director: Patricia Dabbs. Washington, DC.

This page is intentionally blank.

GLOSSARY OF ACRONYMS

Acronym	Gloss
AB	Assembly Bill
ADAD	Assessment Development and Administration Division
AMARD	Accountability, Measurement, and Reporting Division
ART	Administration and Registration Tools
ASL	American Sign Language
CAAs	California Alternate Assessments
CAASPP	California Assessment of Student Performance and Progress
CAPA	California Alternate Performance Assessment
CAHSEE	California High School Exit Examination
CALPADS	California Longitudinal Pupil Achievement Data System
CAT	Computer Adaptive Testing
CCCs	Common Core Connectors
CCSS	Common Core State Standards
CDE	California Department of Education
CDS	County-District-School. Each county, district, and school is assigned a unique CDS code.
CMA	California Modified Assessment
CST	California Standards Test
CTEAG	Connecticut Enhanced Assessment Grant
EADMS	Educators Assessment Data Management System
EC	California *Education Code*
EL	English learners
ELA	English Language Arts/literacy
ETS	Educational Testing Service

FRPM	Free or Reduced Price Meals
FSO	Facility Security Officer
FTP	File Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
HumRRO	Human Resources Research Organization
IAB	Interim Assessment Blocks
ICA	Interim Comprehensive Assessments
IEP	Individualized Education Program
ISAAP	Individual Student Assessment Accessibility Profile
IT	Information Technology
LEA	Local Educational Agency
LEARN	Local Educational Agency Research Network
MAP	Mathematics Assessment Project
MAP	Measures of Academic Progress
MARS	Mathematics Assessment Resource Service
MI	Measurement Incorporated
NCSC	National Center and State Collaborative
NGSS	Next Generation Science Standards
NSLP	National School Lunch Program
NWEA	Northwest Evaluation Association
ORS	Online Reporting System
PII	Personally-Identifiable Information
POC	Point of Contact
RLA	Reading/Language Arts
Smarter Balanced	Smarter Balanced Assessment Consortium
SBE	State Board of Education

SCOE	Sacramento County Office of Education
SD	Standard Deviation
SFTP	Secure File Transfer Protocol
SSID	Statewide Student Identifier
SSL	Secure Sockets Layer
STAR	Standardized Testing and Reporting Program
STS	Standards-based Tests in Spanish
SWDs	Students with Disabilities
TA	Test Administrator
TAG	Technical Advisory Group
TOMS	Test Operations Management Systems
URL	Uniform Resource Locator
USB	Universal Serial Bus

This page is intentionally blank.