TECHNICAL REPORT

Evaluation of the High-Quality Instructional Materials Implementation Grant Program

Prepared by Education Development Center, Inc. for Massachusetts Department of Elementary and Secondary Education.

Lead authors: Sania Zaidi, Neil Schiavo, Anne Wang, Craig Hoyle, Alemayehu Bekele, and Jackie DeLisi

dese MASSACHUSETTS Department of Elementary and Secondary Education

EDC.ORG

Table of Contents

Technical Report Overview

The report provides a detailed description of the methods and instruments that Education Development Center (EDC) used for the evaluation of the Evaluation of the Massachusetts Department of Elementary and Secondary Education (DESE) High-Quality Instructional Materials (HQIM) Implementation Grant program. This report is accompanied with a *summative* report, which presents the findings that answer the evaluation questions and includes discussion and conclusions based on those findings.

The evaluation adopted an *exploratory design* (Newcomer et al., 2015) that was focused on measuring intermediate outcomes to examine the theory of change at the heart of the grant program in order to support future testing in a longer-term study of student outcomes. As an exploratory design, the evaluation was designed to build an understanding of the different contexts in which implementation occurred, with attention to examining the factors that could support or limit success, and to test the influence of experiences on a selection of educator outcomes theorized as indicators of implementation.

Therefore, the goal of this design was to understand the hypothesized model for supporting HQIM implementation, thereby enhancing the understanding of the components in MA DESE's theory of change. EDC's approach included two phases of data collection:

- Phase 1 focused on collecting formative information to guide MA DESE's oversight of the overall grant program and to inform the work of the implementation consultants (ICs) contracted to work directly with grantee districts.
- Phase 2 included analyses to report on findings related to the implementation indicators to document districts' progress in implementing HQIM and to examine related outcomes.

The grant program involved a variety of participants from grantee districts across the state. The evaluation design included data collection activities to represent the perspectives of multiple groups, as summarized in Table 1.

Table 1. Summary of data collection methods

Stakeholder Group	Data Collection Method
Educators in HQIM classrooms	• Teacher survey • Classroom observations • Interviews
Principals and other school-level administrators	• District Implementation Grant Leadership (DIGL) survey • Document review
District-level administrators	• DIGL survey • Document review
Implementation consultants	• Implementation consultant perspectives survey (ICPS) • Interviews • Document review

Evaluation Questions

The evaluation was designed to answer a set of questions, organized into different phases:

- Short-term questions focused on providing formative feedback based on initial grantee activities.
- Mid-term questions focused on proximal indicators of HQIM implementation.
- Long-term questions explored relationships between the proximal indicators and changes in implementation.

Table 2 lists the final set of evaluation questions, which were originally derived from the request for responses and proposal and then updated during the evaluation in response to shifts in grant program implementation.

Table 2. Evaluation questions and data sources

Evaluation Questions	IP Review	IC Interview (Time 1)	IC Interview (Time 2)	Teacher Survey (Time 1)	Teacher Survey (Time 2)	DIGL Team Survey	IC Perspec-tives Survey	Class-room Observa-tions	Teacher Interviews	FOI Matrix
SHORT-TERM RESEARCH QUESTIONS (SRQ)										
SRQ1. What is the composition of District Implementation Grant Leadership (DIGL) teams?						X				
a) What work-roles are represented on grantee DIGL teams?						X				
b) What are the demographics of grantee DIGL teams?						X				
SRQ2. What are the key components of districts' HQIM implementation plans?	X									
a) To what extent do district plans include components or "enabling conditions" for successful HQIM implementation supported by literature?	X									
b) How and to what extent is racial equity centered in district plans?	X									
MID-TERM RESEARCH QUESTIONS (MRQ)										
MRQ1. What current district and school systems and structures (e.g., systems and structures to support			X			X	X			X

coaching, collaboration, professional learning) around the use of HQIM, data collection, and student support are evident, and vary by district characteristics?										
a) What are the perspectives of teachers and DIGL team members on how systems and structures support progress toward meeting their goals?				X	X	X			X	
MRQ2. What are tangible ways in which teachers and districts are centering culturally and linguistically sustaining practices in their implementation of HQIM?			X	X	X				X	
MRQ3. What current educator beliefs and practices (e.g., educator mindsets about instructional materials, curriculum literacy, pedagogical practices, and expectations of students) are evident?			X	X	X			X	X	
LONG-TERM RESEARCH QUESTIONS (LRQ)										
LRQ1. What changes in teacher attitudes, practices, and beliefs, and district and school systems and structures are evident after one year of FC185 support?				X	X		X			X

a) What are the perceptions of DIGL team members and teachers regarding the association between components of HQIM implementation and these changes?						X				

Description of Participants

The evaluation data collection strategy was aligned to the scope of the initiative and its goal of influencing district-level implementation in the grantee sites. Data sources included roles that were closely connected with the articulation of an implementation strategy in each district:

- **District Implementation Grant Leadership (DIGL) team.** Each grantee district created an implementation team as part of the grant requirements. Team composition varied by district but typically included district- and school-administrators, content area coaches and/or coordinators, and classroom teachers.
- **Implementation Consultants.** Each district was assigned one of nine implementation consultants (ICs) to support them during the grant period. The Implementation consultant organizations had a dedicated team that worked with the districts and the schools.
- **Classroom teachers and other school-based educators.** All teachers implementing the HQIM in the 52 grantee districts were part of the teacher population.
- **School-level administrators.** Building and school level administrators, such as school principals, were also included in the study.

The data collection strategy was shaped by facets of the initiative and other considerations set by MA DESE. MA DESE was mindful of limiting the demands placed on school and district staff and limiting the number of groups communicating with district coordinators. The Evaluation Team did not communicate directly with teachers and had limited communication with district and school administrators. Instead, the Evaluation Team worked closely with the ICs in several aspects of data collection. The ICs completed detailed district inventories that listed the membership of each district DIGL team, as well as the number of teachers engaged in the implementation of the HQIM. These inventories were the basis of the Evaluation Team's understanding of the size of the population of these groups. The ICs also supported the administration of the teacher and DIGL surveys. For the teacher survey, the ICs communicated and coordinated with district and school leaders to ensure that all teachers (1) received the survey link to complete the survey, (2) understood the importance of the survey, and (3) and understood why it was being shared with them.

The Evaluation Team derived an understanding of the populations of the different groups involved in the grant program through the district inventories and grantee applications submitted to MA DESE. Table 3 lists the data collection activities organized by data source

and includes a description of each group; the number in each population; and the data collection period (for some data sources, data was collected at two periods of time).

Table 3. Summative HQIM evaluation data collection activities

Data Source	Population	Number	Collection Period	Description	Analysis Procedures
Teacher survey	All teachers implementing the HQIM in the grantee district	4,000 (approx.)	Oct.–Nov. 2023 April–May 2024	Rating scale and open-ended questions designed to identify teacher mindsets, pedagogical practices, and attitudes related to the implementation of the HQIM in their schools and classrooms, including information about teachers, their students, and the PL activities teachers received.	Data were combined with district-level characteristics using hierarchical linear models (HLM) to identify factors associated with changes in mindset and practice and attitudes related to the implementation grant. Open-ended questions at the end of both the Time 1 and Time 2 surveys were qualitatively coded to provide an additional means of surfacing teachers' perspectives related to implementation.
DIGL survey	All members of the DIGL team in each grantee district; all lead school administrators in participating schools	253	May–June 2024	Rating scale and open-ended questions designed to assess the composition and processes of DIGL teams and how teams incorporate racial equity and inclusivity into their district plans to ensure that the implementation of HQIM is equitable and meets the diverse needs of students.	Descriptive statistics were run to summarize the distribution of DIGLs across various demographic categories (e.g., roles, ethnicity, and districts). Weighted means and aggregated measures (combining "strongly agree" and "agree" responses into a single measure) were used to provide a picture of stakeholder engagement and perceptions. Open-ended questions were qualitatively coded to gain insights into DIGL team members' perspectives on HQIM implementation.
ICPS	Each grantee district completed by each IC	52	May 2024	Rating scale and open-ended questions designed to inform each district's progress toward meeting indicators of high-quality implementation.	Weighted means and aggregated measures were used to provide a picture of district progress toward meeting indicators of high-quality implementation. Open-ended questions were coded to give a single rating for communication and for stakeholder engagement.
Classroom observations and	Purposeful sample of classrooms and schools	18	March–May 2024	A sample of classrooms using the HQIM were observed by two researchers who took field	Notes and transcribed interviews were analyzed for themes related to centering the "why" of learning, classroom

teacher interviews				notes during the observation. The teacher of the observed classroom was then interviewed within 1–2 days of the observation about what was observed and the teacher's experiences with the HQIM implementation.	discourse, role of students in learning, role of reflection in learning, and partnership with families in learning.
IC interviews	Each IC organization	9	Sept. 2023 May 2024	Each IC team was interviewed about the role it played in supporting districts' development and the use of a landscape analysis and an implementation plan, as well as support with professional development (PD) and culturally and linguistically sustaining practices (CLSP) (Time 1); reflection on some preliminary findings from the evaluation related to pacing, PD, and scaffolding; and any changes since the last time the ICs were interviewed regarding CLSP (Time 2).	Interviews were transcribed and analyzed for themes related to IC roles in supporting HQIM implementation and district practices and barriers related to pacing, PD, scaffolding, and CLSP.
Landscape analysis review	Completed by each district with their IC	52	June–July 2023	A checklist was developed and used to assess each landscape analysis for elements related to high-quality curriculum implementation. The checklist was then used to create a feedback summary for each district.	Feedback summaries were analyzed for common themes related to instructional vision, HQIM selection, curriculum-embedded PL, leadership support, systems and structures, racial equity, and teacher buy-in.
Implement ation plan review	Completed by each district with their IC	52	Oct.–Dec. 2023	A rubric was developed and used to assess implementation plans for alignment to required elements of the implementation plan (e.g., vision, theory of change), alignment to recommendations from the landscape analysis, and for elements related to high-	The presence or absence of elements in implementation plans was calculated for each component and subcomponent. Components included leadership readiness and responsibility, PD, systems and structures, and communication and stakeholder engagement.

				quality implementation of curriculum.	

Evaluation Instrumentation

The remainder of this technical report provides detailed descriptions of each data collection method employed during the evaluation, with a separate section for each method. All methods are aligned to the HQIM Literature review, which was conducted as an initial step in the evaluation as a theoretical foundation of the research on HQIM implementation. A description of the literature review is provided here as context for the data collection methods that follow. The sections for each method include information about the development of instruments, validation steps (if applicable), the sample and administration process, data management and cleaning, and analysis procedures. Additional information and the instruments are included as appendices to this report.

High-Quality Instructional Materials Implementation Framework

The HQIM Implementation Framework guided the evaluation by identifying the evidence-based elements of HQIM implementation to examine in our evaluation. For example, questions in the teacher survey, DIGL survey, ICPS, teacher interviews, IC interviews, and elements of the fidelity of information (FOI) matrix all align with the framework. In addition, the rubrics to assess district landscape analyses and implementation plans were also grounded in the framework.

The framework is based on a review of the existing research related to implementation. The Evaluation Team conducted a literature search to identify factors related to successful implementation of innovations and curricula.

Developed with rounds of feedback from MA DESE, the framework is aligned with MA DESE's IMplementMA framework, the Coherence Guidebook, and Standards of Effective Administrative Leadership Practice. It is organized with expected outcomes for each of the four IMplementMA phases: Learn & Prepare, Investigate & Select, Launch, and Implement & Monitor. Each outcome has action steps aligned with it. In addition, we created a theory of change to help simplify the framework by identifying responsibilities for each set of actors in the system. It is organized with expected outcomes for each of the four IMplementMA phases: Learn & Prepare, Investigate & Select, Launch, and Implement & Monitor. Each outcome has action steps aligned with them. In addition, we created a theory of change to help simplify the framework by identifying responsibilities for each set of actors in the system.

Fidelity of Implementation Matrix

Based on the framework, the Evaluation Team created a set of indicators, compiled into a matrix, to organize descriptive information about each district's progress toward implementing HQIM. The FOI matrix is a tool that defines these indicators and the processes for determining descriptive ratings of each district's progress toward establishing research-based practices for high-quality implementation. The FOI matrix was developed and enriched through multiple rounds of feedback and revision, which included feedback from MA DESE staff members as well as members from EDC's partner Education Resource Strategies.

The FOI matrix is located in Instrument 8. It contains a description of the four HQIM implementation components and the sets of indicators that describe each district's progress in supporting and implementing HQIM. The matrix outlines a set of 14 prioritized indicators organized by and aligned with the four components within the framework for high-quality implementation of HQIM. Each indicator is listed in a row with a short descriptive title, a definition, currently proposed data sources, and potential ratings. For each indicator, the Evaluation Team defined a set of data sources and steps for analysis and interpretation. The ratings are based on data collected through other methods employed by the Evaluation Team (the teacher survey, District Implementation Grant Leadership team survey, and Implementation Consultant Perspectives survey). In some cases, ratings are based on item scores, while in others, they are based on scale scores. Ratings for most indicators rely on multiple data sources as verification or triangulation of the rating.

By applying the FOI matrix to the data sources, the Evaluation Team produced a spreadsheet that captured the data collected for each district, the team's ratings for each district on each indicator, and the overall ratings for each district for each component. Further analysis included exploratory and descriptive analysis of ratings by district characteristics (e.g., size, urbanicity, student population) and HQIM characteristics (e.g., IC, curriculum, grade band).

Using the *FOI Indicator and Ratings* document as a guide for analysis, the Evaluation Team created a FOI spreadsheet with the resultant set of implementation ratings for each district. In the spreadsheet, the four components are reflected across the top, the indicators in the second row, and within indicator the accompanying data source is labeled in the third row. There are a total of three data sources: the DIGL survey, teacher survey, and ICPS survey.

Following the analysis plan laid out in the FOI indicators document, in most cases, the team averaged across items and across data sources to determine a rating. They used SPSS and/or STATA to write syntax to rename the variables, write question labels, reverse

code (where necessary), and generate descriptive variables to analyze the data. The team manually entered data into the FOI spreadsheet for all indicators and components. Final ratings were determined using the guidance in the FOI matrix to assign a rating between "0" (representing the lowest level of evidence of implementation) to a "3" (representing the highest level of evidence of implementation) for each indicator. Rating assignments were done by pairs of Evaluation Team members with an additional review of all ratings by senior staff to ensure consistency and accuracy.

The Evaluation Team conducted correlational analyses to explore possible relationships between continuous variables among student demographics, and the normalized scores overall and for each component for all districts. These analyses found that variables such as percentage of students who were high needs, ELs, and Black or African American and Hispanic or Latino only yielded coefficients that were not statistically significant, indicating that these variables did not produce notable linear trends.

Teacher Survey – Time 1 and Time 2

Teacher surveys conducted in fall 2023 and spring 2024 were designed to identify teacher mindsets, pedagogical practices, and attitudes related to the implementation of the HQIM grant program in their schools and classrooms. The teacher surveys collected information about teachers, their students, and the professional learning activities in which teachers participated. This information was combined with district-level characteristics to identify factors associated with changes in mindset and practice and attitudes related to the implementation grant. Open-ended questions at the end of both the Time 1 and Time 2 surveys provided an additional means of surfacing teachers' perspectives related to implementation.

The development of the Time 1 and Time 2 teacher surveys began by operationalizing evaluation questions related to investigating teacher mindset and practices. An initial step was the identification of areas of interest to MA DESE and consistent with the literature on HQIM implementation. Seven areas of interest were identified as outcomes.

1. Teacher satisfaction
2. High expectations
3. Perceived equity
4. Use of explicit instruction practices consistent with HQIM
5. Teacher beliefs about students
6. Teacher buy-in

7. Culturally and linguistically sustaining practices

Each area of interest is described in more detail below. Items measuring each of these outcomes were not specifically designed as scales. Nonetheless, the extent to which conceptually similar items fit together to form robust outcomes was examined using factor analysis, and examining the reliability of grouped items based upon theory. With alphas ranging from .702 to .931, the combined outcome scores were used in subsequent multi-level models.

Table 4 shows the number of items used and reliability (Cronbach's alpha) for outcome scores for time 1 and time 2.

Table 4. Reliability of Time 1 and Time 2 outcome scores

Construct	Short Description	Number of Items	Alpha Fall	Alpha Spring
Teacher satisfaction	Satisfaction with elements of HQIM implementation	10	0.931	0.926
High expectations	Teachers report high expectation for all student groups	5	0.702	0.759
Teacher perceived equity	Teacher perception of providing fair and appropriate access to rigorous learning opportunities to all students	9	0.926	0.904
Classroom practices math	Use of appropriate explicit instruction practices while teaching math	7	0.867	0.811
Classroom practices English Language Arts (ELA)	Use of appropriate explicit instruction practices while teaching ELA	9	0.852	0.798
Teacher beliefs about students	Teacher perceptions about students' academic capabilities	7	0.790	0.846
Teacher buy-in	Teacher's support of HQIM and its underlying principles	1 primary	N/A	N/A

Teacher Satisfaction

The teacher satisfaction scale consisted of items designed to measure educators' contentment with various aspects of the curriculum implementation. This scale includes ten items, each focusing on a different facet of the intervention, such as the overall implementation, quality, and ease of use of the curriculum, as well as the effectiveness of workshops and trainings. It also assesses support from principals, feedback on curriculum use, teacher collaboration, student engagement, the curriculum's potential to reduce performance gaps between racial groups, and communication about the implementation's purpose. With a high internal consistency (alpha = .926), this scale allows for an overall assessment of teacher satisfaction with the intervention. Table 5 shows the ten items included in the satisfaction scale. The table also shows the mean satisfaction score without each item included, the correlation of each item to the other items in the scale, and the impact of the item on the scale alpha. All ten satisfaction items strongly correlate with the total score, and positively contribute to the alpha of the scale.

Table 5. Time 2: Teacher satisfaction items

N of items=10, Cronbach's Alpha= .926	Scale Mean if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
How satisfied are you with the following? - a) The implementation overall	37.81	0.811	0.914
How satisfied are you with the following? - b) The quality of the curriculum	37.76	0.802	0.914
How satisfied are you with the following? - c) The ease of using the curriculum	38.11	0.742	0.917
How satisfied are you with the following? - d) Workshops or trainings focused on the curriculum	38.03	0.649	0.922
How satisfied are you with the following? - e) Support from your principal to use the curriculum	37.65	0.567	0.926
How satisfied are you with the following? - f) Feedback on your use of the curriculum	37.92	0.713	0.918
How satisfied are you with the following? - g) Teacher collaboration focused on implementing the curriculum	37.54	0.593	0.924

How satisfied are you with the following? - h) Student engagement with the lessons and units	38.00	0.761	0.916
How satisfied are you with the following? - i) The extent to which the curriculum will reduce performance gaps between different racial groups	38.42	0.750	0.916
How satisfied are you with the following? - j) Communication about the purpose of implementation of the curriculum	37.89	0.770	0.915

High Expectations

The High Expectations scale measures teachers' expectations for students in the HQIM intervention with four items, each focused on different aspects of curriculum effectiveness and student success. The items assess teachers' beliefs about the curriculum's role in supporting high academic standards as outlined by the Massachusetts Curriculum Frameworks, the alignment of these standards with student success, and their confidence in their students' potential for success in school. Additionally, the scale evaluates teachers' perceptions of how the implementation of HQIM might impact the number of students meeting grade-level expectations. With an alpha of .759, this scale demonstrated a moderate level of internal consistency, providing a reliable measure of teachers' expectations suitable for group-level assessment. Table 6 shows the four teacher survey items included in the High Expectations scale and the impact of each on the overall scale. While the extent to which teachers agree that their students will be successful in school is positively correlated with the total score of the other items, it is worth noting that the alpha of this scale would increase if the item were omitted. Given the positive correlation, and the conceptual link of the item with high expectations, it was decided to keep the item.

Table 6. Time 2: High expectations items

N of items=4, Cronbach's Alpha=.759	Scale Mean if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
The curriculum plays a vital role in supporting the implementation of high academic standards of the MA Frameworks.	11.80	0.708	0.635

The curriculum reflects high academic standards of the MA Frameworks are the best way to ensure student success.	12.00	0.729	0.620
To what extent to you agree with the following statement: My students will be successful in school.	11.41	0.343	0.814
To what extent, if any, do you believe the implementation will change the overall number of students in your target class who would meet grade-level expectations.	14.09	0.537	0.733

Perceived Equity

The Perceived Equity scale measured teacher self-reports of the extent to which all students are able to access the rigorous academic content of the HQIM being implemented. The scale consists of twelve items, each addressing different aspects of curriculum inclusivity and support. The items evaluate teachers' agreement on whether the curriculum materials are representative of their students and supportive of English learners, students with IEPs or 504 plans, and those performing below grade-level. Additionally, the scale assesses teachers' beliefs about the equitable treatment of students by teachers and administrators, and the curriculum's attention to students' cultural and linguistic backgrounds. It also measures teachers' perceptions of how the implementation might impact the number of students from various groups (low-income households, English learners, students with IEPs or 504 plans, Black or African-American, and Hispanic or Latino students) meeting grade-level expectations. With an alpha of .904, this scale demonstrated high internal consistency, providing a reliable measure of teachers' perceptions of equity and inclusivity within the curriculum overall. Table 7 shows the twelve items included on the Perceived Equity scale along with their relative contribution to the scale. It is worth noting that the two items asking teachers about their perception of how others treated their students were less correlated to the total score than other items.

Table 7. Time 2: Perceived equity items

N of items=12, Cronbach's Alpha=.904	Scale Mean if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted

To what extent do you agree with the following about the curriculum materials? - d) The images, names, and activities in the curriculum are representative of my students.	34.44	0.581	0.898
To what extent do you agree with the following about the curriculum materials? - e) lessons and activities are supportive of English learners (ELs).	35.17	0.712	0.891
To what extent do you agree with the following about the curriculum materials? - f) The lessons and activities are supportive of students with IEPs and/or 504 plans.	35.34	0.721	0.890
To what extent do you agree with the following about the curriculum materials? - g) The curriculum materials support students who are currently performing below grade-level.	35.48	0.723	0.890
To what extent do you hold the following beliefs about your students in the target class? - e) My students are treated equally and fairly by teachers.	33.27	0.248	0.909
To what extent do you hold the following beliefs about your students in the target class? - f) My students are treated equally and fairly by administrators.	33.49	0.299	0.909
To what extent do you hold the following beliefs about your students in the target class? - g) The curriculum attends to my students' cultural & linguistic backgrounds and experiences.	34.43	0.643	0.894
To what extent, if any, do you believe the implementation will change the number of students in your target class who would meet grade-level expectations for the following groups: Students from low-income households	36.55	0.728	0.891
To what extent, if any, do you believe the implementation will change the number of students in your target class who would meet grade-level expectations for the following groups: Students who are English Language Learners	36.68	0.736	0.891

To what extent, if any, do you believe the implementation will change the number of students in your target class who would meet grade-level expectations for the following groups: Students with IEPs and/or 504 plans	36.67	0.737	0.891
To what extent, if any, do you believe the implementation will change the number of students in your target class who would meet grade-level expectations for the following groups: Students who are Black/African American	36.47	0.713	0.891
To what extent, if any, do you believe the implementation will change the number of students in your target class who would meet grade-level expectations for the following groups: Students who are Hispanic or Latino	36.50	0.727	0.890

Math Explicit Instruction

The Math Explicit Instruction practices scale measured teachers' use of a set of instruction strategies when using the implemented HQIM. The scale consists of seven teacher practices. These items assess the frequency with which teachers focus on various instructional aspects in their target class, including mathematical procedures and algorithms, problem-solving strategies, mathematical reasoning, conceptual understanding, developing mathematical models, constructing mathematical arguments, and making mathematical connections. With an alpha of .811, this scale demonstrated good internal consistency, indicating that it reliably measures the extent to which teachers employ explicit instruction practices in their math teaching consistent with HQIM adoption. Table 8 shows the seven classroom practices included in the Math Explicit Instruction scale. Correlations of individual items with the scale score with that item removed were positive ranging from .37 to .61, and all items contributed to the alpha of the scale.

Table 8. Time 2: Math explicit instruction items

N of items=7, Cronbach's Alpha=.811	Scale Mean if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
How often do you focus on these aspects of instruction in the target class where you	27.77	0.370	0.788

use the curriculum? - a) Mathematical procedures and algorithms			
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - b) Problem-solving strategies	27.44	0.556	0.746
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - c) Mathematical reasoning	27.40	0.616	0.742
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - d) Conceptual understanding	27.47	0.600	0.738
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - e) Developing mathematical models	27.72	0.540	0.742
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - f) Constructing mathematical arguments	27.98	0.524	0.755
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - g) Mathematical connections	27.63	0.539	0.743

English Language Arts (ELA) Explicit Instruction

The ELA Explicit Instruction practices scale measured teachers' use of a set of instruction strategies specific to this content area when using the implemented HQIM and consists of nine items. These items assess the frequency with which teachers focus on various instructional aspects in their target class, including reading foundational skills, vocabulary and knowledge development, active engagement with complex texts, text-specific tasks and assignments, active discussion with complex texts, writing conventions and sentence structures, the writing process, writing craft, and language development for English learners. With an alpha of .798, this scale demonstrated good internal consistency, indicating that it reliably measures the extent to which teachers employ common explicit instruction practices in ELA that are consistent with HQIM practice. Table 9 shows the nine ELA practices comprising the ELA Explicit Instruction scale. The results show that teacher

reports of individual practices were strongly related with all items contributing positively to the scale alpha.

Table 9. Time 2: ELA explicit instruction items

N of items=9, Cronbach's Alpha=.798	Scale Mean if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - a) Reading foundational skills	35.14	0.431	0.767
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - b) Vocabulary and knowledge development	34.57	0.430	0.750
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - c) Active engagement with complex texts	34.52	0.409	0.753
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - d) Text-specific tasks, questions, and assignments with complex texts	34.56	0.370	0.756
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - e) Active discussion with complex texts	34.50	0.367	0.758
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - f) Writing conventions and sentence structures	35.00	0.669	0.710
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - g) Writing process	35.05	0.583	0.725

How often do you focus on these aspects of instruction in the target class where you use the curriculum? - h) Writing craft	35.11	0.543	0.730
How often do you focus on these aspects of instruction in the target class where you use the curriculum? - i) Language development for English learners	35.20	0.504	0.743

Beliefs about Students

The Beliefs about Students scale measured teachers' beliefs about student agency and attitudes towards learning in the implemented HQIM. The scale consists of eight items. These items assess teachers' perceptions of their students' ability to find help at school, learn challenging material, value their schoolwork, and improve academic abilities through effort. Additionally, the scale evaluates beliefs about the equal and fair treatment of students by teachers and administrators, the curriculum's attention to students' cultural and linguistic backgrounds, and the overall success of students in school. With an alpha of .846, this scale demonstrated high internal consistency, indicating reliable measurement of teachers' beliefs about their students. This scale is distinct from the High Expectations scale, which focused on teachers' expectations for student success and the impact of curriculum standards on student achievement. In contrast, the Beliefs about Students scale emphasized teacher beliefs about student support, equity, and the role of the curriculum in addressing diverse student needs with higher scores indicating conditions consistent with students developing a growth mindset. Table 10 shows the eight items included in the scale, and that responses to all the included items were consistent for this group of teachers.

Table 10. Time 2 Teacher beliefs about students items

N of items=8, Cronbach's Alpha=.846	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
To what extent do you hold the following beliefs about your students in the target class? - a) My students are able to find help at school when having difficulties.	19.157	0.578	0.817
To what extent do you hold the following beliefs about your students in the target class?	18.896	0.631	0.810

- b) My students have the ability to learn challenging material.			
To what extent do you hold the following beliefs about your students in the target class? - c) My students value their school work.	18.322	0.639	0.808
To what extent do you hold the following beliefs about your students in the target class? - d) My students' academic abilities will increase through effort.	20.292	0.551	0.821
To what extent do you hold the following beliefs about your students in the target class? - e) My students are treated equally and fairly by teachers.	20.578	0.554	0.821
To what extent do you hold the following beliefs about your students in the target class? - f) My students are treated equally and fairly by administrators.	19.694	0.500	0.827
To what extent do you hold the following beliefs about your students in the target class? - g) The curriculum attends to my students' cultural & linguistic backgrounds and experiences.	18.423	0.482	0.836
To what extent do you hold the following beliefs about your students in the target class? - h) My students will be successful in school.	19.253	0.687	0.805

Teacher Buy-In

While the teacher survey did not directly include a Teacher Buy In scale with respect to the intervention, the item “On a scale from 1-10, to what degree has the HQIM curriculum influenced your instruction in the target class?” serves as a proxy for teacher buy-in by quantifying the extent to which teachers perceive the HQIM intervention as impactful on their teaching practices. A higher score indicates a greater perceived influence, suggesting stronger acceptance and integration of the HQIM curriculum into their instruction. This metric helps gauge overall teacher engagement and willingness to adopt the new curriculum and remains highly correlated with the items used for the satisfaction index. Its correlation with the satisfaction index was .804.

Culturally and Linguistically Sustaining Practices

While many teacher survey items relate to culturally and linguistically sustaining practices (CLSP), a set of items were included specifically form this scale. Teachers were asked about their professional learning experiences both prior to and during the intervention, specifically regarding these practices. The survey also gauged their interest in further training and whether they felt such training was unnecessary. Additionally, it assessed the impact of professional learning on their ability to develop culturally relevant instructional activities, the level of support from their principal, and the frequency of coaching events related to culturally and linguistically sustaining practices. Table 11 lists the CLSP items included on the Time 2 survey.

Table 11. Items Addressing culturally and linguistically sustaining practices

Included in Time 1 and Time 2 Surveys
I received professional learning prior to the intervention.
I received training on culturally and linguistically sustaining practices during the intervention.
I have not received specific professional learning about culturally and linguistically sustaining practices but am interested.
I have not received specific training, but culturally and linguistically sustaining teaching is not an area I need to work on.
In general, the professional learning experiences offered by my school help me develop culturally relevant instructional activities.
Only Included in Time 2
During the past year, to what extent has the administration advocated for or supported culturally and linguistically sustaining practices?
Based on your own experience, during the 2023-2024 school year, how frequently did coaches focus on culturally and linguistically sustaining practices during these meetings?

Survey items were developed through a combination of adapting existing items from released surveys and constructing items based upon specifics of the HQIM implementation. An initial item bank with more than 200 candidate items was constructed from these items (Kane et al., 2016; Martin et al., 2013; Mullis et al., 2019; NTPS, 2021; TALIS, 2018). The final survey consisted of approximately 100 multiple-choice items and took approximately 15 minutes to complete (see Instrument 1 for copy of Time 2 teacher surveys). Items included demographic questions about teachers, their classroom composition, classroom practices, expectations, buy-in with respect to HQIM, attitudes toward equity, satisfaction with the implementation of the HQIM, as well as experiences with professional learning. While the original survey design did not call for the creation of specific scale scores, it did provide multiple measures of each of the constructs of interest. Table 12 lists the mid-term research questions #1, #2, and #3, as well as long-term

research question #1. The table also lists the key outcomes, as well as the analytic methods and instruments used to answer these research questions.

Table 12. Teacher surveys: Key outcomes, categories, and analysis methods
Note: "HLM" stands for "hierarchical linear models."

Evaluation Question	Instrument	Key Outcomes	Breakout Categories	Methods
MRQ1. What current district and school systems and structures (e.g., systems and structures to support coaching, collaboration, and professional learning) around the use of HQIM, data collection, and student support are evident and vary by district characteristics?	Time 2 teacher survey	• PL activities • Administrative support • Assessments • Curriculum materials	• District characteristics	• Descriptive statistics • 2-level HLMs (and districts)
MRQ2. What are tangible ways in which teachers and districts are centering culturally and linguistically sustaining practices in their implementation of HQIM?	Time 2 teacher survey	• Culturally and linguistically sustaining practices	• Teacher years of experience • Type of Curriculum • Class size • District characteristics • PL	• Descriptive statistics
MRQ3. What current educator beliefs and practices (e.g., educator mindsets about instructional materials, pedagogical practices, and expectations of students) are evident?	Time 2 teacher survey	• High expectations • Student Beliefs • Teacher buy-in • Explicit instruction • Perceived Equity • Teacher satisfaction	• Teacher years of experience • Type of Curriculum • Class size • District characteristics, • Professional learning (PL)	• Descriptive statistics • 2-level HLMs (teachers, and districts)
LRQ1. What changes in teacher attitudes, practices, and beliefs, and district and school systems and structures are evident after one year of FC185 support?	Time 1 and Time 2 matched sample	• PL activities • Administrative support • Assessments • Curriculum materials	• District characteristics	• Descriptive statistics • 2-level HLMs (teachers, schools, and districts)

Changes to the Time 2 survey included the addition of items to inform the FOI matrix. Nineteen items were added to the Time 2 survey. These additions included questions that asked teachers for their perspectives on leadership's role in supporting implementation, teacher collaboration time, instruction time for HQIM lessons, use and availability of curriculum materials, student support services, and two-way communication with school administration. To keep administration time comparable to Time 1, teacher practice variables not considered directly associated with HQIM were removed from the survey.

The open-ended question at the end of the Time 2 survey was also revised to ask teachers to reflect on the extent to which they believed the implementation has changed since they first started working with the curriculum. In the Time 1 survey, the question was left open so that teachers could identify comments on topics of their choosing.

Instrument Validation

Multiple validation steps were taken at different points in the survey development and review of collected data. Validating the initial survey instrument involved conducting a small piloting process with critical reviewers prior to the survey launch, examining school-level characteristics of responding and non-responding schools using publicly available data, conducting both confirmatory and exploratory factor analyses on survey items, and examining alphas for items designed to measure the same construct.

A pilot with critical reviewers was conducted prior to administering the survey to teachers in the field. A total of 10 participants were recruited: five participants were EDC technical assistant staff (not part of the Evaluation Team) with experience working with teachers and survey design, and five teachers not directly connected with the schools served by the grant. These teachers were recruited informally through existing networks with a priority on teachers from Massachusetts.

The pilot was structured such that reviewers took the survey first, followed by a structured interview conducted by Evaluation Team with the purpose of testing key assumptions about the survey. Of particular interest was the length of time to complete the survey; the appropriateness of language and terms; and survey functionality, design, and flow.

EDC staff were asked to provide overall thoughts and feedback and to comment on or identify the following:

- Items that would be unclear or otherwise raise concerns
- The extent to which items addressing racial equity were appropriate

- Overall length of the survey
- The extent to which questions would capture teacher attitudes, beliefs, and practices with respect to the implementation of HQIM in their schools

Teachers were asked similar questions as well as questions about the professional development (PD) items, pedagogical content knowledge, and the extent to which they believed teachers would be able to provide accurate information about racial equity.

The pilot process resulted in a set of recommendations, such as editing the language of item stems for greater clarity, editing for a better flow for the survey taker, and using consistent terminology. These recommendations were incorporated into the final instrument.

Sample Analysis

While every effort was made to obtain survey responses from all participating schools, there were a number of instances where schools, for whatever reason, did not respond to the survey. To determine the extent to which non-response at the school-level might impact results, publicly available data provided by the National Center for Education Statistics was compared for both responding and non-responding schools within each district. The factors examined for sample screening were local, Title 1 status, as well as the type of curriculum being implemented. Table 13a-c shows that number and percentage of schools whose teachers responded to the surveys in the Fall of 2023 and the Spring of 2024.

Table 13a. Percentage of schools responding

Percentage of Schools Responding	
Fall 2023	Spring 2024
86%	88%

Table 13a shows that, at the school-level, a similar percentage of schools responded to the survey in Fall of 2023 and Spring of 2024 with 14% having no respondents in Fall of 2023 and 12% having no respondents in 2024.

Table 13b. Percentage of total schools responding and not responding: Math and ELA

Type of Curriculum	Schools Responding				Schools Not Responding			
	Fall 2023 n=197		Spring 2024 n=200		Fall 2023 n=31		Spring 2024 n=28	
	n	pct.	n	pct.	n	pct.	n	pct.
Math	127	64.5	128	64	22	71	21	75
ELA	70	35.5	72	36	9	29	7	25

Not all schools were matched with NCES data. Missing data is not included in these percentages.

Table 13b. shows that 64.5 percent of the schools responding to the 2023 survey and 64% of the schools responding to the 2024 survey reported being math teachers. Among schools not responding in each of the administrations of the survey the pattern remained similar with 71 percent of schools not responding administering a math curriculum in 2023 and 75 percent administering a math curriculum in 2024.

Table 13c. Percentage of total schools responding and not responding: Title 1 status

Title I Schools	Schools Responding				Schools Not Responding			
	Fall 2023		Spring 2024		Fall 2023		Spring 2024	
	n	pct.	n	pct.	n	pct.	n	pct.
Title I Status (Yes)	133	67.5	131	65.5	14	45.2	16	57.1

Not all schools were matched with NCES data. Missing data is not included in these percentages.

Table 13c. shows fall and spring school-level responses for Title 1 schools. A similar percentage of schools responding to the survey were Title 1 schools in both administrations. Among non-responding schools, a higher percentage of Title 1 was observed (57.1%) in the spring administration.

Table 13d. Percentage of total schools responding and not responding: School Location

School Location	Schools Responding				Schools Not Responding			
	Fall 2023 n=193		Spring 2024 n=94		Fall 2023 n=27		Spring 2024 n=26	
	n	pct.	n	pct.	n	pct.	n	pct.
Urban	7	3.6	8	4.1	3	11.1	2	7.7
Suburban	164	85	165	85.1	21	77.8	20	76.9
Town	6	3.1	6	3.1	-	-	-	-
Rural	16	8.3	15	7.3	3	11.1	4	14.3

Not all schools were matched with NCES data. Missing data is not included in these percentages.

Table 13d. shows that the percentage of schools responding by location was similar in both administrations of the survey. Most teachers responding were from suburban schools in both administrations.

Factor Analysis

While items included to measure each construct were not chosen specifically to form scales, items were tested to see how well they fit together by construct using data from the first administration of the survey. Confirmatory factor analysis (CFA) provided support for the 7-factor model, where math and ELA explicit instruction are treated separately. Exploratory factor analysis (EFA) results, on the other hand, supported a single dominant factor suggesting that among this population of teachers, the constructs were highly related. This same picture was supported when looking at alphas for combinations of items where items expected to group together fit with strong item-total correlation; when tested, combining all scale score items produced a solid alpha of .95. Alphas on the potential scales were similar from Fall to Spring and ranged from .70 for the combined 5 items looking specifically at high expectations, to .93 for the 10-item satisfaction scale.

Four different CFA models were fitted to the data and an EFA was conducted. All models were estimated using full information maximum likelihood (FIML) to keep as many cases as possible. Alpha reliability and omega reliability were estimated to determine internal consistency among items after an EFA had been conducted. In terms of the 7-factor model the root mean square error of approximation (RMSEA) indicates an acceptable level of model misspecification. The standardized root mean square residual (SRMR) indicates decent model fit to the data. However, the Tucker-Lewis Index (TLI) indicates poor model fit with a value of 0.75 well below the .95 cutoff. The RMSEA and the SRMR estimates are just below their respective acceptable cutoffs and the TLI is well below the .95 cutoff indicating good model fit. This model does not have great fit to the data. An examination of the standardized loadings shows that there are 21 items with loadings less than 0.40. This indicates a weak relationship between the item and the latent constructs. Modifying the model by removing these items did improve model fit.

Based upon earlier EFA results, two one-factor CFA models were also conducted. The first with all items and the second with the same items that were included in the modified model. The fit of both models was worse than the initial 7-factor CFA. Table 14 shows the model fit indices for the seven factor and one factor CFA models. Due to smaller sample sizes, EFA and Math were considered in separate CFA models.

Table 14. Model fit indices for CFA models

CFA Model	Chi Square (DF)	P value	RMSEA	90% CI		SRMR	TLI
				LL	UL		

7 Factor	30944.21 (2316)	<.001	0.07	0.07	0.07	0.07	0.75
7 Factor (mod.)	20363.57 (1248)	<.001	0.15	0.15	0.15	0.10	0.61
1 Factor	61491.06 (2345)	<.001	0.10	0.10	0.10	0.10	0.48
1 Factor (mod.)	33095.95 (527)	<.001	0.15	0.15	0.15	0.10	0.61
Cutoffs			<0.08			<0.08	>0.95

Models for ELA and math were run separately as only a subset of teachers answered each section of the survey. Table 15 shows model fit information for the ELA sample. The initial 7-factor confirmatory factor analysis (CFA) model for the ELA sample demonstrated better overall model fit to the data compared to the single factor model in terms of fit indices RMSEA and the SRMR.

Table 15. Model fit indices for ELA CFA models

CFA Model	Chi Square (DF)	P value	RMSEA	90% CI		SRMR	TLI
				LL	UL		
7 Factor	10221.62 (2316)	<0.001	0.07	0.07	0.07	0.07	0.74
7 Factor (mod.)	8872.92 (1986)	<0.001	0.07	0.07	0.07	0.07	0.76
1 Factor	18327.78 (2345)	<0.001	0.10	0.10	0.10	0.10	0.47
1 Factor (mod.)	9012.84 (594)	<0.001	0.14	0.14	0.15	0.09	0.60
Cutoffs			<0.08			<0.08	>0.95

Table 16 shows model fit information for the math sample. Like results based on the ELA sample, the initial CFA model for the math sample demonstrated better overall fit to the data compared to the single factor model in terms of the RMSEA and the SRMR.

Table 16. Model fit indices for Math CFA models

CFA Model	Chi Square (DF)	P value	RMSEA	90% CI		SRMR	TLI
				LL	UL		
7 Factor	16196.33 (2316)	<0.001	0.07	0.07	0.07	0.07	0.74
7 Factor (mod.)	12324.85 (1683)	<0.001	0.07	0.07	0.07	0.06	0.79
1 Factor	30583.97 (2345)	<0.001	0.10	0.10	0.10	0.10	0.48
1 Factor (mod.)	16300.87 (560)	<0.001	0.15	0.14	0.15	0.10	0.60
Cutoffs			<0.08			<0.08	>0.95

The results remain consistent with the expectation that it would be hard to cleanly separate these constructs in this implementation.

Sampling and Administration

The identification of the target sample of teachers within each school was initially done by having ICs complete a "district inventory." The inventory included fields where the consultant confirmed the HQIM curriculum and grade in each school, where the curriculum was taught, and the number of teachers using the HQIM in that school. These inventories were then crosschecked with information provided by DESE and compared against school size and the expected number of teachers. Where potential discrepancies were noted, attritional information was requested from the IC.

In both administrations of the teacher survey, the inventory information was used to track responses and determine response rates by school. As it was not feasible to identify teachers individually, surveys were administered through an online platform (Qualtrics) via anonymous link. ICs served as contacts with their respective districts and, where possible, encouraged districts to administer the survey during common meeting times for teachers, such as faculty or grade-team meetings. During administration, the Evaluation team provided ICs with weekly response rates at the school- and district-levels to facilitate

follow up with their points of contact and encourage participation. To further increase the likelihood of school and teacher responses, the evaluation provided a modest financial incentive paid to each grantee district that submitted required information. The incentive funds were intended to be spent on refreshments or food for those teachers who completed the survey or to organize an event where teachers would complete the survey.

The use of an anonymous link and its broad applications led to the potential of teachers and other staff not directly involved with the HQIM curriculum taking the survey. Questions confirming the respondents' roles, and the curriculum taught became necessary screens on the survey. Due to the use of an anonymous survey link, teachers were also asked to provide self-generated identifiers. Tables 17a, 17b, and 17c show teacher demographics, average class size and years of teaching experience.

Table 17a. Teacher demographics

Teacher Demographics			
Teacher Race/Ethnicity	Time 1 (n=1791)	Time 2 (n=1400)	Matched Sample (n=502)
Black/African American	2.46%	2.03%	1.9%
Asian	1.07%	1.31%	1.7%
Hispanic	2.00%	2.18%	2.3
White	74.95%	83.97%	88.2
Native American	0.46%	0.51%	0.6
Native Hawaiian, Pacific Islander	0.19%	0.07%	0
Multi Race, Non Hispanic	0.70%	0.51%	0.6
Prefer not to respond	9.57%	12.11%	8.8

Table 17b. Average class size

Average Class Size	
Time 1 (mean, *SD*)	20.83 students, *4.43*
Time 2 (mean, *SD*)	20.97 students, *4.10*
Matched Sample (mean, *SD*)	20.87 students, 6.63

Table 17c. Years of teaching experience

Years of Teaching Experience	
Time 1 (mean, *SD*)	14.95 years, *9.20*
Time 2 (mean, *SD*)	15.22 years, *9.21*

Matched Sample (mean, *SD*)	15.07 years, 9.33

Data Cleaning

Anonymous links were used for survey responses, requiring the first step in cleaning to be ensuring that the target population was represented and properly identified for analysis. Respondents from each school were compared with the data collected by ICs on the number of teachers and the curriculum being used in each district and school. Global screening rules based upon survey items were applied to ensure that teachers responding to the survey were grant participants working with the chosen HQIM in the correct schools and districts. Educators who led classroom instruction were included in analyses (i.e., respondents who identified as support staff or aides were excluded).

Where teachers answered unexpectedly (e.g., a respondent selected a different HQIM than the one listed for their district), results were reviewed on a case-by-case basis. In some cases, where teachers identified using a different curriculum, it was determined that a valid response was entered. A teacher was kept in the sample if data were consistent with participation at that school; otherwise, they were flagged for removal. A separate inclusion flag was constructed for teachers who had provided complete self-generated identifiers.

Analysis Procedures

Two analytical samples from teacher survey data were used to answer evaluation questions about teacher's beliefs, attitudes, and practices related to the HQIM implementation. Time 2 teacher survey data were used to answer questions about which key teacher outcomes (high expectations, beliefs about students, classroom practices, focus on equity, teacher buy-in, and teacher satisfaction) were evident. A smaller sample of matched teachers for Time 1 and Time 2 using self-generated IDs was used to examine the extent to which these factors changed over time (from Fall to Spring).

For both sets of analyses, two-level multi-level models (MLMs) were employed to look at teacher outcomes in order to account for the nested structure of the data. At Level 1, the model looked at teacher outcomes. At Level 2, districts' data were analyzed by including district scores from the FOI analyses.

Initially, unconditional models were estimated for each outcome variable to decompose the total variance into components attributable to teachers and districts. This step was done to determine the proportion of variance at each level, indicating the extent of variability that could be explained by district characteristics. Models then incorporated predictor variables at each level for each of the key teacher outcomes. (e.g., teacher demographics and experience at Level 1, and district policies and funding at Level 3 to identify factors contributing to the observed outcomes). When assessing change from Time 1 to Time 2, a variable representing survey administration was included at the teacher level.

Outcome variables (scales) examined using multi-level models:

- High expectations
- Beliefs about students
- Classroom practices (explicit instruction)
- Beliefs about equity
- Teacher buy-in
- Teacher satisfaction with the implementation

Variables included at Level 1 (Teacher):

- Teacher years of experience
- Curriculum
- Class size
- Percentage of students who are designated as English learners (ELs)
- Percentage students with Individualized Education Plan or 504-plan
- Percentage of students from low-income households
- Ethnicity of the teacher
- Years prior experience with the curriculum
- Working with other teachers to review student data
- Taking part in professional learning
- Attitude toward professional learning
- Experience with culturally and linguistically sustaining practices

Variables included at Level 2 (District)

- District level leadership readiness and responsibility
- District level support for professional development
- District level systems and structures
- District level communication and stakeholder engagement
- District level years of implementation experience with the curriculum
- District level percentage of high needs students
- District level percentage of low-income students

The two –level models examining spring teacher survey (Time 2) outcomes takes the following form:

$$
\begin{aligned}
Y_{ij} = & (\gamma_{00} + \gamma_{01} W_{1j} + \gamma_{02} W_{2j} + \ldots) \\
& + (\gamma_{10} + \gamma_{11} W_{1j} + \gamma_{12} W_{2j} + \ldots) X_{1ij} \\
& + (\gamma_{20} + \gamma_{21} W_{1j} + \gamma_{22} W_{2j} + \ldots) X_{2ij} \\
& + \ldots \\
& + (\gamma_{p0} + \gamma_{p1} W_{1j} + \gamma_{p2} W_{2j} + \ldots) X_{pij} \\
& + e_{ij}
\end{aligned}
$$

Where:

- **Teacher-Level Predictors (*XpijX_{pij}*Xpij)**: These variables capture characteristics and experiences specific to individual teachers, such as years of teaching experience, prior training, and classroom practices. Each predictor *XpijX_{pij}*Xpij has an associated effect size (*γp0\gamma_{p0}*γp0) that describes its impact on the outcome.
- **District-Level Predictors (*WqjW_{qj}*Wqj)**: These variables describe broader contextual factors at the district level, such as leadership, professional development opportunities, and demographic characteristics. Each district-level predictor influences the intercept (*γ00\gamma_{00}*γ00) and the slopes (*γ10,γ20,...\gamma_{10}, \gamma_{20}, \ldots*γ10 ,γ20 ,...) of teacher-level predictors, representing how district conditions modify the relationship between teacher characteristics and the outcome.
- **Intercept and Slopes**: The model includes an intercept (*γ00\gamma_{00}*γ00) representing the baseline outcome level across all districts and slopes (*γp0\gamma_{p0}*γp0) representing the influence of each teacher-level predictor on the outcome. District-level variables (*WqjW_{qj}*Wqj) adjust these baseline levels and slopes to capture differences across districts.
- **Error Term (*eije_{ij}*eij)**: This component captures the variation in the outcome that is not explained by the predictors included in the model, representing individual differences among teachers.

See Appendix A for the individual model results for Time 2 outcomes.

HLM model specification for LRQ 1

As part of LRQ1, the evaluation team examined changes in teacher satisfaction, expectations, equity in HQIM implementation, and perceptions of student capabilities. This analysis compares matched teacher data collected at two periods—Time 1 (Fall) and Time 2 (Spring)—to assess changes in teacher attitudes, practices, and beliefs. The dataset consists of 1,006 observations from teachers who participated in the survey at both time points, matched using self-generated IDs.

Statistical assumptions checks:

- **Normality**: The normality of key outcome variables was verified using box plots and histograms, which indicated that most variables were roughly symmetric with few outliers and minor deviations such as slight skewness, suggesting that no transformations were necessary.
- **Multicollinearity**: Variance Inflation Factors (VIFs) were calculated for all predictors, with values remaining below 5, confirming that multicollinearity did not compromise the integrity of the analyses, and thus all predictors were retained in the models.

Model specifications:

- Using Unconditional Models, the evaluation team decomposed the total variance into components attributable to individual teachers (Level 1) and districts (Level 2), establishing the baseline variance without introducing predictor variables.

- Using Conditional Models, we estimated with the introduction of predictor variables at both the teacher (Level 1) and district (Level 2) levels. Using these models, the evaluation team explored how various factors influence observed changes from Time 1 to Time 2.
    - **Level 1:** The conditional models incorporated several teacher-specific variables (see the previous section for the full list of variables).
    - **Level 2 :** The only district-level variable included in the analysis is the district identifier.

Each model is structured as follows:

Model 1: Teacher satisfaction with HQIM Implementation

$y_{ijk}=\gamma_{00}+\gamma_{01}(\text{Timeijk})+\gamma_{02}(\text{Prior experience}_{ijk})+\gamma_{03}(\text{Class size}_{ijk})+\gamma_{04}(\text{Non-White}_{ijk})+\gamma_{05}(\text{Percentage of ELL in class}_{ijk})+\gamma_{06}(\text{Percentage of IEP in class}_{ijk})+\gamma_{07}(\text{Percentage of low-income students in class}_{ijk})+\gamma 08(\text{Teachers high expectations}_{ijk})+\gamma_{09}(\text{PL positive impact on HQIM Instruction}_{ijk})+\ldots+u_{0j}+v_{0k}+e_{ijky}$

Model 2: Teachers' high expectations

$y_{ijk}=\gamma_{00}+\gamma_{01}(Time_{ijk})+\gamma_{02}(Prior\ experience_{ijk})+\gamma_{03}(Class\ size_{ijk})+\gamma_{04}(Belief\ about\ student\ capabilities_{ijk})+\gamma_{05}(Non\ White_{ijk})+\gamma_{06}(Percentage\ of\ ELL\ in\ class_{ijk})+...+u_{0j}+v_{0k}+e_{ijky}$

Model 3: Teacher perceived equity in HQIM Implementation

$y_{ijk}=\gamma_{00}+\gamma_{01}(Time_{ijk})+\gamma_{02}(Prior\ experience_{ijk})+\gamma_{03}(Class\ size_{ijk})+\gamma_{04}(PL\ meeting\ teacher\ needs_{ijk})+...+u_{0j}+v_{0k}+e_{ijky}$

Model 4: Teacher belief in student capabilities

$y_{ijk}=\gamma_{00}+\gamma_{01}(Time_{ijk})+\gamma_{02}(Prior\ experience_{ijk})+\gamma03(Class\ size_{ijk})+\gamma_{04}(Relevant\ PL\ for\ HQIM_{ijk})+...+u_{0j}+v_{0k}+e_{ijky}$

Where:

- y_{ijk} : Outcome variable for the *i*-th teacher in the *j*-th district and *k*-th match ID.
- γ_{00} : Intercept, the overall average effect when all predictors are at their reference levels.
- $\gamma_{01}, \gamma_{02},..., \gamma_{09}, ..., \gamma_{09}$: Fixed effects coefficients for the predictors.
- $Time_{ijk}$: time point (e.g., Fall or Spring).
- $Prior\ experience_{ijk}$: Years of experience.
- $Class\ size_{ijk}$: Number of students in the teacher's class.
- $Non\text{-}White_{ijk}$: Indicator variable for non-White teachers.
- $Percentage\ of\ ELL\ in\ class_{ijk}$: Proportion of English Language Learners in the class.
- $Percentage\ of\ IEP\ in\ class_{ijk}$: Proportion of students with Individualized Education Programs.
- $Percentage\ of\ low\text{-}income\ students\ in\ class_{ijk}$: Proportion of low-income students.
- u_{0ju} : Random effect for district *j*.
- v_{0k}: Random effect for match ID *k*.
- e_{ijk}: Residual error term for the *i*-th teacher in the *j*-th district and *k*-th match ID.

The results of each of the above models are detailed in Appendix B.

Qualitative Analysis of Open-Ended Responses

For Time 1, a sample of 685 open-ended teacher survey responses was coded by a five-person qualitative data analysis team. Codes for the codebook were created using a combination of deductive code creation—using the HQIM evaluation document created by the team for review of landscape analyses—and inductive code creation—based on themes emerging from the open teacher survey responses. The resultant codebook

contained 213 codes, as seen in Instrument 11 (Dedoose Codes for Teacher Survey). Approximately 550 responses were independently coded by the research team before reaching the point of saturation and repetition of codes. Each coder completed their coding and then met weekly to discuss emergent codes. Four hundred responses were dual coded for inter-rater agreement.

Coding was facilitated through the use of Dedoose, a qualitative coding software. Dedoose was chosen due to the data analysis team's familiarity with the software, as well as the ability of Dedoose to allow multiple coders to access the data simultaneously and its data analysis tools. Prior to coding, the team received an hour-long training on Dedoose. The team met weekly throughout the coding process to discuss pertinent findings from their individual coding assignments.

Given the size of the codebook, the "code presence" function in Dedoose was useful in helping coders understand the distribution and co-mapping of codes and in supporting the data analysis process. Data were grouped into similar code areas and presented to the coding team for discussion before higher-level themes were drafted.

For Time 2, the second administration of the teacher survey yielded a sample of 937 open-ended responses for another round of coding. In Time 2, one team member completed an overview of all the responses and then selected 10 responses for the team to code and to facilitate a discussion of the range of responses to the new open-ended question: *Thinking about the {HQIM} implementation over time, do you feel the implementation in your classroom has changed since you've started using this curriculum? In what ways?*

The four-person coding team received individual coding assignments and met weekly to discuss emergent findings. The team used the existing codebook from Time 1, which was supplemented by memos, to share thoughts on how the coded segment(s) answered the new open-ended question. Three new codes were added to the codebook: "No change," "Examples of changes over time," and "Not sure." Two hundred codes were dual coded to ensure inter-rater reliability.

District Implementation Grant Leadership (DIGL) Team Survey

The data from the DIGL survey played an important role in evaluating the effectiveness of DIGL teams as they implemented HQIM in their districts. By examining the composition and processes of these teams, the survey reports on their structures and implications for their strategies for implementation. In addition, the survey data provided insights into how teams incorporate racial equity and inclusivity into their district plans, aiming to ensure that the implementation of HQIM is equitable and meets the diverse needs of students. One of the primary purposes of the survey was to inform the development of the FOI matrix

by incorporating the perspectives and roles of DIGLs in the implementation of HQIM. This inclusion was important for tailoring the FOI matrix to accurately reflect and support the strategic efforts of leadership teams in improving the implementation of HQIM.

The development of the DIGL survey initially considered the use of focus groups to collect nuanced insights from DIGL teams. However, due to logistical challenges and the specific timing required for data collection, the approach was amended to use a survey format instead. This decision was driven by the need for a broader, more efficient method for collecting data that could also support the development of the FOI matrix and ratings.

Building on this framework, the Evaluation Team crafted survey questions applicable to all DIGL team members, with additional targeted questions for school-level administrators. Aligning the DIGL survey with the FOI matrix involved using its components and indicators as a starting point, thereby ensuring that the survey questions directly addressed the critical aspect of HQIM implementation. Table 18 displays a summary of the constructs in the DIGL survey and accompanying analysis procedures, organized by evaluation question.

Table18. Summary of survey constructions, analysis procedures, and evaluation questions

Evaluation Questions	Survey Construct	Analytic approach
What is the composition of District Implementation Grant leadership (DIGL) teams? • What roles/representation on DIGLs are necessary for effective implementation plan development? • What processes, structures, and resources (e.g., IC services) are in place to support implementation plan development? • Which processes, structures, and resources (e.g., IC services) best support effective implementation plan development?	DIGL demographics: • Involvement with DIGL team activities, curriculum use, professional development, and professional learning Leadership: • Instructional vision, building/school administrator role and administrators provide instructional support for educators	• Use frequencies and percentages to analyze demographic characteristics. • Simplify data interpretation by combining similar response categories (“strongly agree” and “agree”) into a unified measure for clearer analysis.
What processes did DIGLs employ to make use of formative information from the evaluation (e.g., continuous improvement data study; stakeholder engagement)? • Did formative information result in improved implementation plans?	Communication and stakeholder engagement: • Two-way communication and involvement of diverse voices and reflection of district demographic in committees	• Used weighted mean percentage for Likert scale responses, distribution counts, and percentages for frequency data. • Simplify data interpretation by combining similar response categories (“strongly agree” and “agree”) into a unified measure for clearer consensus analysis.
What changes in district and school systems and structures are evident	Systems and structures:	• Use weighted mean percentage for Likert scale responses,

after one year of HQIM implementation? • What are the perceptions of DIGL team members and teachers regarding the association between components of HQIM implementation and these changes?	• Resource availability, DIGL team functionality and support and continuous improvement and data use Professional development: • Collaboration frequency and adequacy as well as coaches/ instructional expert support	distribution counts and percentages for frequency data. • Simplify data interpretation by combining similar response categories ("strongly agree" and "agree") into a unified measure for clearer consensus analysis. • Analyze professional learning and collaboration using frequencies, percentages, and cross-tabulation.

Sampling and Administration

The initial sample for the DIGL survey was derived from the district inventory data, which included DIGL members from 52 districts. To refine our target group and to ensure a focused approach, we used additional data from grant applications (including fields for DIGL team membership and emails for those staff), which resulted in a list of 238 potential DIGL team members. One early challenge was the variation in the designations of DIGL team members across different districts, such as "DIGL team," "implementation committee," or "team." In addition, the composition of these teams varied, with some districts having a predominantly teacher-based team and others consisting mainly of building leaders.

To create consistency in terminology and to ensure accurate identification, we developed the following working definition of *DIGL team members* in consultation with MA DESE:

> *DIGL team members* are individuals directly involved in the leadership, coordination, implementation, and support of the HQIM initiative. This typically includes a small team composed of a combination of district and building-level administrators, curriculum leads, and instructional coaches responsible for overseeing curriculum implementation, professional development, and grant activities.

This definition was reviewed and affirmed by the ICs and the Evaluation Team. To further validate and finalize our list, the Evaluation Team also contacted each IC responsible for their respective districts and asked them to verify the accuracy of the DIGL contact information the team had compiled. For those few instances where the team did not have accurate email addresses, they checked school and district websites to add principal

emails. After validating this list with all nine ICs, they distributed the DIGL survey to a total of 253 members.

The evaluation received a total of 167 completed DIGL surveys for a response rate of 66%, with a representative from nearly all districts (51 of 52 districts represented). Respondents represented a range of job roles (see Table 19). Descriptive analyses were conducted to examine the ethnic demographics of respondents (see Table 20) and their years of work experience (see Table 21).

Table 19. DIGL respondents by role

Role	Frequency	Percent
Building-level administrator (e.g., principal, assistant principal, deans)	38	23%
Building-level curriculum lead (e.g., curriculum coordinator, specialist for specific grades/subjects)	24	14%
District-level administrator (e.g., superintendent, chief academic officer)	56	34%
Instructional coach (e.g., literacy coach, math coach)	35	21%
Lead or co-lead teacher	23	14%
Total	167	100%

Table 20. DIGL respondents by race/ethnicity

Race/Ethnicity	Frequency	Percent
Asian	3	2%
Black or African American	1	1%
Hispanic	2	1%
Multi-Race, Non-Hispanic	1	1%
Prefer not to respond	12	7%
White	136	81%
Native American	1	1%
Non-respondents	11	7%
Total	167	100%

Table 21. DIGL respondents by years of experience

Overall Years of Experience	Frequency	Percent

11-25 years	88	61%
5-10 years	5	3%
Less than 5 years	1	1%
More than 25 years	50	35%
Total	144	100%

Data Cleaning

The analytical procedures for the DIGL survey data began with a comprehensive data cleaning process. The raw survey data was downloaded from Qualtrics and then uploaded to Box, a secure cloud-based platform, for collaborative work. For the first step in cleaning the data, the Evaluation Team identified and removed survey responses that were test responses (from internal quality assurance testing), duplicate responses, and responses that were more than 60% incomplete. Once these data were clean, the team used SPSS and/or STATA to write syntax to rename the variables, write question labels, reverse code (where necessary), and generate descriptive variables to analyze the data.

In addition, this step included checking for missing data, correcting any outliers, and ensuring consistency in response categories. These activities maintained the quality of the analysis, ensuring that findings would be reliable and accurately reflect true data patterns. Following this step, an internal feedback cycle was conducted with the internal DIGL survey team to identify any potential data issues and verify that the analysis aligned with the overall objectives of the DIGL survey and broader evaluation goals.

Analysis Procedures

Once the data were prepared, the Evaluation Team implemented its analysis plan, utilizing a mixed-methods approach that combined both quantitative and qualitative techniques. Quantitative analysis included applying descriptive statistics, such as frequencies, percentages, and cross-tabulations, to summarize the distribution of DIGLs across various demographic categories (e.g., roles, ethnicity, and districts). This analysis enabled the team to identify patterns and trends in the data, particularly in areas such as involvement with DIGL team activities, curriculum use, and the frequency of professional learning activities. Weighted means and aggregated measures (combining “strongly agree” and “agree” responses into a single measure) were used to assess consensus levels across different survey items, providing a clear picture of stakeholder engagement and perceptions.

Open-Ended Responses Analysis Steps

The Evaluation Team reviewed the 55 open-ended responses it received, focusing on identifying key words and phrases to gain insights into DIGL team members’ perspectives

on HQIM implementation. This work was conducted by two members of the Evaluation Team who reviewed each response. This process allowed the team to extract important themes relevant to structures and supports (such as professional development and coaching), reports on changes over time and teacher buy-in.

Implementation Consultants Perspectives Survey

The purpose of the implementation consultant perspectives survey (ICPS) was to provide data on each grantee district to contribute to the FOI analysis. For several FOI indicators, this data served the important role of triangulating data from teacher and DIGL team member perspectives. The Evaluation Team's goal was to gather data about specific aspects of the implementation, including the following:

- Alignment of the districts' instructional vision to the HQIM being implemented
- Role of building administrators in implementation
- Teachers' use of curriculum materials
- Time for teacher collaboration
- Frequency of DIGL team meetings
- Opportunities for families and other stakeholders to be involved in implementation efforts
- Whether or not stakeholder groups represented district demographics

The ICPS addressed the following evaluation question, "What changes in district and school systems and structures are evident after one year of HQIM implementation?"

The Evaluation Team used data from the ICPS to inform ratings regarding each district's progress toward meeting pre-determined indicators of high-quality implementation.

Development Process

ICPS items were developed in parallel with DIGL survey items and teacher survey Time 2 items related to implementation. As the Evaluation Team developed the FOI indicators, they considered which indicators could be best informed by IC perspectives and triangulate the perspectives of DIGL team members and teachers. The team developed the items to align directly with the FOI indicators, which were reviewed by MA DESE leadership to ensure validity and feasibility of response. Items were first provided in a Google document for ICs to review and then later entered into a Qualtrics survey, which was administered to ICs. See Table 22 for a detailed list of the ICPS constructs and survey items.

Table 22. ICPS constructs and associated items

Construct	Items
Leadership	To what extent do you believe the district's instructional vision is aligned with the expectations of the HQIM?
Leadership	Please describe any evidence you have that the district's instructional vision is regularly used to guide HQIM implementation.
Leadership	To what extent are building administrators involved in any of the following: • Attending or supervising classroom support • Aligning school-based systems and structures to HQIM expectations • Ensuring racial equity is centered in HQIM implementation
Professional development and professional learning	Please rate the extent to which you agree with the following statement and then comment on your ratings in the space below: • There is sufficient time set aside in the teachers' schedules for teachers to collaborate and plan for using HQIM with fidelity.
Systems and structures	Please rate the extent to which you agree with each of the following statements and then comment on your ratings in the space below: • All teachers are using the curriculum materials with fidelity to the curriculum expectations and pacing.
Systems and structures	Please rate the extent to which you agree with each of the following statement and then comment on your ratings in the space below: • Teachers in this district regularly supplement the curriculum with additional materials or lessons.
Systems and structures	Please rate the extent to which you agree with each of the following statement and then comment on your ratings in the space below: • The District Implementation Team meeting frequency is sufficient to support HQIM implementation.
Communication and stakeholder engagement	What opportunities do community or family stakeholders have to be involved in HQIM implementation?
Communication and stakeholder engagement	In what ways do the stakeholder groups and committees represent the demographics of the district?

Sampling and Administration

The surveyed population was the entire group of ICs. Each IC organization was directed to fill out the survey once for each of the districts they worked with, resulting in a single response per district. The survey was completed for all 52 districts, for a response rate of 100%.

Data Cleaning

IC data were reviewed for completion and consistency in responses. There were seven instances where an IC submitted two surveys for a single district. For these instances, an EDC researcher examined these data to determine which responses to include in the final dataset. In most of these instances, it was evident which survey was intended to be used (e.g., only one question was answered for one survey while all questions were answered for the other survey). In these cases, complete survey responses were retained. In the two instances with double data, the ratings were averaged. In one case, the IC answered all questions except the open-ended response questions and indicated in the second survey that the Evaluation Team should use the response from the first survey for those questions. In another instance, there was a missing response, but an associated open-ended response question was answered. In that case, we reviewed the description and assigned a rating.

For one question, the Likert scale on the survey questions was accidentally reversed, resulting in Qualtrics assigning the highest rating (5) to the lowest ratings (1) and vice-versa. One member of the Evaluation Team reverse coded the responses to ensure the highest score of five was consistent with the meanings.

There were two open-ended response questions about communication and stakeholder engagement. These items were given a numeral rating. To score them, two members of the Evaluation Team reviewed these questions separately and each assigned a rating. Where the team members disagreed, they came to a consensus about what the rating should be.

Analysis Procedures

ICPS data were used to inform FOI ratings, which described districts' progress toward implementing essential components of HQIM. For each district, ICPS data provided one data point that informed ratings.

Implementation Consultant Interviews

The IC interviews allowed the Evaluation Team to understand the role of the ICs and to dig deeper into patterns observed by the Team in the evaluation data. Since each consultant organization worked with multiple districts and divided their districts across their staff, the consultants were interviewed as a group, with the intention that all members of each organization involved in the HQIM Implementation grant program would be present at the interview. This allowed for the interviews to capture variation in the responses that reflected the varied experiences in implementation across the grantee districts.

The first set of interviews examined the role that ICs played in supporting districts' development and use of the landscape analysis and implementation plan as well as support with PD and culturally and linguistically sustaining practices (CLSP). The second set of interviews asked districts to reflect on some preliminary findings from the evaluation related to pacing, PD, and scaffolding, as well as any changes since the last time the ICs were interviewed regarding CLSP. Table 23 provides brief summaries of the topics that were included in each interview protocol for Time 1 and Time 2.

Table 23. ICs group interview topics

Time 1	Time 2
Development and use of landscape analysis	Systems and structures: Instructional time
Development of implementation plans	Systems and structures: Professional development
Selection of professional development	Systems and structures: Scaffolding
Partnerships with pre-service programs	Centering of CLSP
Factors that have led to success so far	
Centering of racial equity	

Sampling

Members of all nine IC groups participated in Round 1 and Round 2 interviews. Participants included project leads as well as IC staff members who worked directly with schools.

Analysis Procedures

In the Time 1 and Time 2 IC interviews, two Evaluation Team members conducted team interviews on Zoom. The nine IC organizations were interviewed separately, and the interviews were recorded. One Evaluation Team member served as a lead interviewer, and the second was the note taker for each interview. Interviews were transcribed using Rev and transcriptions were stored in Box. The Evaluation Team reviewed the transcriptions for accuracy, using the recording to correct any errors in the transcripts. Below, the processes for the analyses of the interview questions are listed for Time 1 and Time 2.

Analysis Process for Implementation Consultant Interviews Data Analysis Procedures

Time 1:

- Week 1: The lead interviewer randomly selected three transcripts for initial examination. Team members read through this initial set of transcripts, noting

interviewees' words that seemed most important or salient for illustrating ICs' role and their work with districts.

At the end of Week 1, the team met to discuss initial impressions, focusing on important points and ideas that were relevant to the questions and that might need to be tracked. The team developed an initial coding scheme that began with each of the team's questions as codes, and based on their review and discussion, they added any additional codes that might help them identify patterns or themes across the interviews.

- Week 2: Team members independently identified text in the three transcripts that addressed each of the agreed-upon codes. Sample text was recorded in a spreadsheet that included a separate tab for each team member. On each tab, rows were the codes/questions with one column per interview.

  At the end of Week 2, the team met again to discuss coding, paying attention to areas of agreement and disagreement. Depending on the level of agreement or disagreement and the number of additional codes, the team conducted one more round of coding, either on the same three transcripts or on one to two additional transcripts.

- Week 4: The team divided into pairs, applying the updated set of codes to the remaining interviews.

- Week 5: Once all interviews were coded using the refined coding scheme, the team compiled coded text into one tab in the spreadsheet. Team members again divided into pairs, with each pair reviewing half of the codes. This review focused on identifying patterns or themes in interviewee responses. The pairs met to discuss themes and recorded each theme, producing a synthesis that summarized the essential points within each code. Team members also identified one to two segments of coded text as examples for each theme.

- Week 6: The lead team member reviewed and compiled the theme syntheses to create an overall summary of interview findings.

Time 2:

- Interview notes were organized by interview questions. A lead and a second interview team member were assigned to identify themes for each interview question.

- Team members independently analyzed each question for themes and met to resolve any differences.

- A summary of themes and the percentage of ICs who exhibited that theme, along with a few quotes to exemplify each theme, were created to be shared with MA DESE.

Teacher Observations and Interviews

The purpose of observing classrooms and interviewing teachers in the classrooms observed was to gather insight into the educator practices (e.g., educator mindsets about instructional materials, pedagogical practices, culturally and linguistically sustaining practices) used by teachers in classrooms where HQIM was being implemented. The Evaluation Team's goal was to analyze the data by employing thematic analysis (Creswell, 2009) to understand and illustrate educator practices in the classroom. The team collected classroom data from multiple grade bands and from both math and English language arts (ELA) HQIMs. In addition to this, the teacher interviews that followed classroom observations provided an opportunity to collect data on the structures and supports available to the teachers, such as the role of administrators and professional learning opportunities the teachers received, the latter specifically in pedagogy around culturally and linguistically sustaining practices. The data from the teacher interviews and observations addressed the following evaluation questions (see Table 24):

Table 24. Teacher interviews and observations by evaluation question

Evaluation Questions	Data Sources	Methods
SRQ1b. What are the perspectives of teachers and DIGL team members on how systems and structures support progress toward meeting their goals?	Teacher interviews	Thematic analysis
MRQ2. What are tangible ways in which districts are centering culturally and linguistically sustaining practices in their implementation of HQIM?	Teacher interviews	Thematic analysis
MRQ3. What current educator beliefs and practices (e.g., educator mindsets about instructional materials, curriculum literacy, pedagogical practices, and expectations of students) are evident?	Classroom observations (field notes) Teacher interviews	Thematic analysis

Development Process

The Evaluation Team drew from culturally relevant pedagogy (Ladson-Billings, 1995) and the framework for social justice education (Hackman, 2005) to inform the constructs of interest. The research indicates that socially just teaching and instruction is associated with increased learning gains, especially for students from nondominant groups. By increasing the relevance of content and the use of culturally responsive teaching, which are a knowledge base for social justice, teacher increase student opportunities for

learning. Culturally responsive teaching has been shown to positively impact students' achievement, participation, motivation, and sense of belonging in educational environments (e.g., Grant & Gillette, 2006; Gay, 2000; Ladson-Billings,1995). Furthermore, it has helped shape identity development and positive attitudes about learning among racial and linguistic minority students (e.g., Rodriguez et al., 2004; Waxman & Tellez, 2002).

In addition to these frameworks, the constructs were informed by the MA DESE culturally and linguistically sustaining practices resource document (see Figure 1), which emphasizes asset-based teaching, high expectations and support, cultural competence and community building, sociopolitical awareness, and partnerships with students and families.

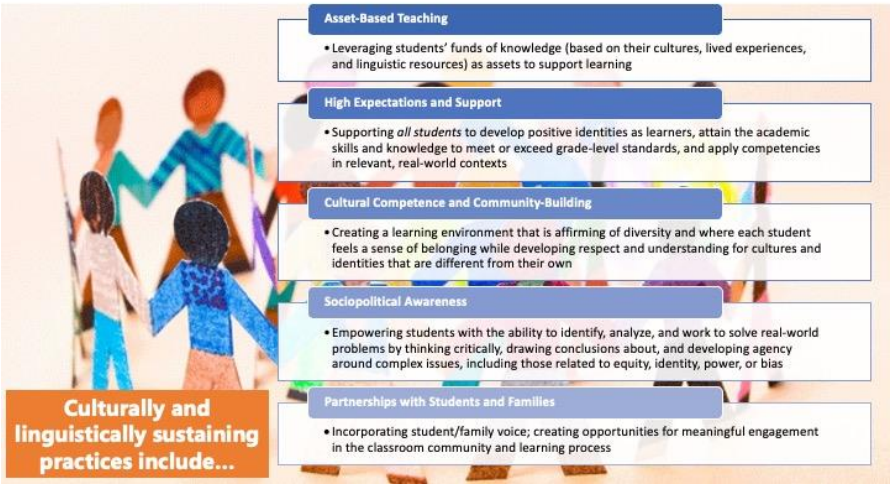


Figure 1. *A resource for culturally and linguistically sustaining practices. Reproduced with permission from MA DESE.*[1]

In total, the Evaluation Team had five overarching constructs of interest:

1. Centering the "why" of learning (e.g., sharing criteria for success, connections to students lives)
2. Classroom discourse (e.g., teacher-student discourse, peer-peer discourse)
3. Role of students in learning (e.g., student agency and voice)
4. Role of reflection in learning (for both teacher and students)
5. Partnership with families in learning (incorporating family voice)

See Table 25 for the full list of constructs.

[1] Massachusetts Department of Elementary and Secondary Education. (2024). Internal Document.

Table 25. Constructs of interest for classroom observations

Constructs	Sub-Constructs	Indicators
Centering the WHY of learning	Sharing learning intentions and criteria for success for students	How and when are learning intentions and criteria for success on them shared in the classroom? We might capture this by observing: • Teacher sharing of the learning goal with students • Driving question written on the board • Making learning relevant to students lives in and/or out of school • Documenting how the criterion for success is shared with students when there is an activity happening • Discussing what success looks like through discussions or reminders about grade-level expectations, stress on terminology, etc.
	Use of instructional materials to promote student learning	What materials are being used and how?
	Use of tools to promote student learning	What tools are being used and how to promote student learning?
	Providing feedback	How is feedback provided to move students forward in their learning?
	Eliciting student thinking	In what ways is student thinking elicited? • Examples of the diversity of ways in which this can be done are written, oral, drawing, storytelling, and through song.
	Promoting student learning	What are the types of supports provided to promote student learning? • Examples of differentiated learning and personalized learning.
	Translanguaging	Are there any instances of translanguaging? • Examples of encouraging students to access their full linguistic repertoires to demonstrate or make connection to student learning.
	Connections to students lives	How is learning being connected to students' lives both in and outside of school? Examples:

		• Cross-disciplinary connections • Connections to pop culture • Connections to current events • Connections to school activities
Classroom discourse	Teacher-student discourse	What is teacher-student discourse most focused on, and when does it happen most frequently? • Problem-solving • Knowledge building • Clarifying instructions for an activity • Consensus building in discussions
	Classroom discussions	What do classroom discussions look like, and what role do they play in sensemaking and knowledge building? • Do discussions decenter whiteness by allowing for nondominant viewpoints, encouragement for a pluralistic view of the world, allowing counter-storytelling, etc.? • How do teachers handle race-based interactions, comments, etc.? • Do discussions provide multiple points of entry into the learning? • Do discussions connect to prior learning? • Do discussions make cross-disciplinary connections? • Do discussions connect to students' lives and interests? • Do discussions surface or address socio-politics topics (e.g., equity, power, identity, bias)?
Role of students	Student agency	What agency do students have in their learning? • Choice of modality when demonstrating thinking (e.g., written, verbal, drawings) • Student choice in forming a small group for discussions • Student encouragement to volunteer answers versus teacher's choosing of students
	Student voice	How is student voice used in the classroom? • Students participate in sense making. • Student ideas are used to build on learning.

	Students as instructional resources for one another	What does student-student interaction look like?
	Translanguaging	What are the examples of translanguaging by students (i.e., use of home language as a vehicle to demonstrate their thinking)?
	Connections to student life	How are students making connections to their lives inside and outside of school?
Reflection	Teacher reflection	What methods are used by the teacher to engage students in reflection? • Teacher looks at student work and realizes that an adjustment needs to be made in instruction • Teacher looks at student work and then adjusts instruction • Teacher talks about changing course of planned instruction
	Student reflection	Are there opportunities for students to reflect on the learning, and if so, how?
	Critical analysis	Is there any critical analysis done of the informational texts? • Discussion on different points of views • Counter storytelling
	Tools for critical analysis	What tools are used for critical analysis during discussions, individual work, or pair-shares?
Partnership with families	Incorporating family voice	Is there evidence of engagement from families in the learning process?

Sampling

The Evaluation Team employed a purposeful sampling technique and identified a total of nine districts for classroom observations. In each district, they identified one school and conducted two classroom observations at the school site. In total, the Evaluation Team conducted 18 classroom observations and teacher interviews.

The sampling criteria employed was as follows:

1. Schools predominantly serving African American/Black and/or Hispanic/Latino students
2. Evidence from teacher survey Time 1 data where respondents indicated a high level of satisfaction (satisfied and very satisfied) in overall implementation of HQIM in their schools
3. Sample included classrooms using ELA and Math HQIM
4. Sample included a range of grade bands K–12
5. Historical information from MA DESE based on its work with school districts

See Table 26 for a detailed description of the set of sample schools.

Table 26. Sample for classroom observation and teacher interview

School [Pseudonym]	Grade	Subject Area	HQIM Used	Implementation Consultant
Butterfly Elementary 2	5	Math	iReady Math	STEM Learning Design
Butterfly Elementary 1	1	Math	iReady Math	STEM Learning Design
Rose Elementary 1	3	Math	Eureka Math	SchoolKit
Rose Elementary 2	5	Math	Eureka Math	SchoolKit
Dogwood High 1	9	Math (Algebra 1)	Agile Mind	UnboundED
Dogwood High 2	11	Math (Algebra 2)	Agile Mind	UnboundED
Lily Middle High 1	7	Math	Eureka Math	SchoolKit
Lily Middle High 2	9	Math (Algebra 1)	Eureka Math	SchoolKit
Iris Elementary 1	1	Math	HMH into Math	Teaching Lab
Iris Elementary 2	3	Math	HMH into Math	Teaching Lab
Orchid Middle 1	6	Math	Illustrative Math	Teaching Lab

Orchid Middle 2	7	Foundational math	Illustrative Math	Teaching Lab
Tulip Elementary 2	5	ELA	HMH into Reading	TNTP
Tulip Elementary 1	6	ELA	HMH into Reading	TNTP
Fern Elementary 1	2	ELA	Wit and Wisdom	Rivet
Fern Elementary 2	3	ELA	Wit and Wisdom	Rivet
Marigold Elementary 1	2	ELA	Wonders	Education First
Marigold Elementary 2	3	ELA	Wonders	Education First

Data Collection and Analysis Procedures

Two observers from the Evaluation Team were present in each classroom observed. The observers used field notes to collect data during classroom observations. Each observation lasted one class period. Across the different sites, the observation time for each classroom ranged from 55 to 90 minutes.

The two observers independently took field notes using a template that divided the class period into 5-minute segments (see the template in Instrument 6 Teacher Observation Field Notes). This allowed observers to capture their observations in 5-minute segments. After the classroom observation, the observers met together to reflect on the observed classroom and consider the constructs of interest. Each observer also individually identified questions to ask the teacher in the interviews. Within 24 hours of the observation, the two observers met with the principal investigator (PI) of the project to debrief about the observation and to collaboratively decide on the questions to ask the teacher in the interview.

Teacher interviews were conducted within 48–72 hours of the classroom observation. Each interview was 30 minutes in length. The interviews were conducted virtually over Zoom and transcribed. The PI and one observer who was at the classroom observation was present at each interview. The Evaluation Team took notes during the interview, and the meeting was also transcribed. The first 20 minutes of the interview were structured around the observed classroom and the last 10 minutes of the interview focused on teachers' questions about their experience with HQIM use and the supports and professional learning they had received.

The interview was transcribed using Rev (a speech-to-text tool) and reviewed manually. The notes and the transcript were manually compiled into one document and served as the data source for the teacher interview.

The data from the classroom observations and the teacher interviews were de-identified. When gathering the data, pseudonyms for the school and teachers were used throughout. File naming conventions also included the use of pseudonyms. Once the observers had taken field notes, they reviewed their notes to remove any identifiable information (e.g., teacher name, student names), and the information from the field notes was entered into a spreadsheet. This spreadsheet then served as a data source for further analysis. The transcripts and field notes were uploaded to a secure cloud storage only after the information had been de-identified. The interview recordings were deleted once the transcripts had been reviewed manually.

The Evaluation Team of five members conducted a thematic analysis following this six-step process (Bruan & Clarke, 2006): (1) familiarization, (2) coding, (3) generating themes, (4) reviewing themes, (5) defining, and (6) naming themes and locating exemplars. The team looked at observation data with corresponding interview data to develop a rich description of instructional practices that were evident in school districts implementing HQIM. The team also did a thorough overview of the data through iterative cycles of reading.

Prior to coding, the team used a deductive approach where they grouped the codes into initial categories that come from their theoretical constructs. Evaluation questions directly informed the development of a robust codebook, which included codes related to topics such as establishing the why of learning (e.g. teacher expectations, setting learning intentions, and teacher beliefs); student engagement; use of materials and instructional supports; culturally responsive and linguistically sustaining environment; classroom discourse (for teachers and students); communication; and reflection.

Next, the team employed open coding followed by the development of focused codes (Charmaz, 2014; Saldaña, 2021). They used the codes to identify patterns and developed candidate themes around educator practices when using HQIM in classrooms and how and when culturally and linguistically sustaining practices were centered. Two team members coded each data set, with one serving as a primary coder and the other as a secondary coder.

EDC's qualitative sub-team met routinely throughout the process to discuss potential new codes, articulate preliminary findings, and discuss emergent themes. To ensure inter-coder reliability within the Evaluation Team's coding process, the sub-team explicitly discussed

variation in coding styles early in the engagement process, iteratively revising the coding to reflect such changes.

After observation data and interview transcripts were coded, the Evaluation Team developed a set of initial themes at the classroom level. They returned to the data and compared these initial themes against the data to ensure that the themes were both useful and an accurate representation of the data. This step led to the construction of themes.

Following this process, the Evaluation Team conducted a cross-case thematic analysis that yielded a synthesis of key findings and insights across all classrooms and schools in the sample. This synthesis complimented other data sources (e.g. teacher surveys) to tell the story of instructional practices and teacher beliefs across the sample of nine schools implementing HQIM.

Teacher Interview. Teacher interview data were organized into a master spreadsheet that provided an overview of the practices discussed in the interview, the challenges reported, and the structures and supports around HQIM that the teachers shared. This organization allowed the team to find patterns in the data and to see what practices related to the constructs that were discussed in each interview (based on their presence in the classroom observation). It also highlighted the challenges reported by the teachers and the types of ongoing professional learning support they had received for HQIM implementation from the district and their building administrators.

Vignettes for Culturally and Linguistically Sustaining Practices. The Evaluation Team used the teacher interview data as well as the classroom observation data to map the instructional practices they had observed in the classroom onto MA DESE's culturally and linguistically sustaining practices (see Figure 1 in prior section). This work focused on answering the evaluation question: MRQ2**.** *What are tangible ways in which districts are centering culturally and linguistically sustaining practices in their implementation of HQIM?*

Two members of the Evaluation Team worked on identifying the observed instructional practices that contributed to unpacking MA DESE's culturally and linguistically sustaining practices. They then identified examples from classrooms that were used to create vignettes illustrating how a practice was used in the classroom.

Classroom Observation. Following are the results from the analysis of the classroom observation data (N = 18 classrooms):

Themes emerged across multiple schools. Following are the top eight themes:

1. Facilitating discourse
2. Students using each other as a resource-collaboration
3. Using materials to promote learning

4. Communication expectations and framing why
5. Connecting learning to students' lives
6. Practicing and reinforcing classroom norms

- Eliciting student thinking through questioning
- Teachers using affirming language

Table 27 summarizes the themes that were observed across all 18 classrooms.

Table 27. Emergent themes around HQIM instruction and learning from classroom observations (n=18)

Thematic Presence	Number of observations in which the theme was present
Facilitating discourse	10
Communication expectations and framing why	8
Using materials to promote learning	6
Students using each other as resource and for collaboration	5
Connecting learning to students' lives	3
Practicing and reinforcing classroom norms and routines	3
Teacher using affirming language	3
Eliciting student thinking through questioning	3
Student agency and voice	2
Students using each other as resource and for collaboration	2
Multiple ways to promote learning	2
Teacher using affirming language	3
Instructional variety	2
Multiple levels of instructional support for scaffolding and supplemental materials	2
Instructional variety	2
Reciprocal conversations	1
Student engagement promoting facilitation of higher-order thinking	1
Using systems and structures to support student learning	1
Modeling activities for students	1

Teacher Interviews

This section covers the results from the analysis of the teacher interview data (N = 18). The Evaluation Team summarized teachers' reports of success in implementing HQIM (Table 27), their challenges with implementing HQIM (Table 28, the supports they received to implement HQIM (Table 29), and the PD they received regarding culturally and linguistically responsive teaching and learning (Table 30).

For the successes reported with HQIM implementation, the most salient ones were *support for collaboration time* (8 out of the 18 observations), *improved student outcomes with HQIM* (8 out of 18), and *increased student engagement* (5 out of 18).

Table 28 details the documented successes, the frequency of occurrence in the sample of 18 teachers, and examples from the teacher interviews.

Table 28. Teacher reported successes with HQIM implementation

Success	Description of Success	Frequency of Occurrence	Example from Teacher Interviews
Support for teacher collaboration time		**8**	
	HQIM team meetings		Curriculum team meetings with whole math team held once a month during PD times.
	Additional collaborative time built in		Leading up to MCAS at the end of the year, 50-60 minutes were built in for teacher collaborative time and to get more support needed to meet the goals of HQIM.
	Grade level meetings		Frequent grade-level math team meetings were held.
	School meetings		School meetings every Friday.
	Planning time		Collaborative planning time was held.
	Collaborative time		Collaborative time with teachers in the district was held.
	Collaborative time		Teacher has collaborative time with two other 6th grade math teachers was held after hours via text or phone or after school in the classroom.
	Weekly collaborative time		Weekly teacher collaborative time was held.
Improved student outcomes		**8**	
	Student agency		Predicting story problems improved students' agency and got students thinking

			differently on what was happening in problems.
	Self-reflection		Teacher modeled self-reflection to students.
	Growth through homework		HQIM assigned one night a week. Teacher thinks students made progress in practicing and solidifying because of homework.
	Growth of tiers 2 and 3 students		Students with specialized learning plans made the most growth last year as indicated by last year's data.
	Student-to-teacher ratio		Having a co-teacher allowed students to have more than one adult to seek as help. Special education teacher was not exclusive just to students with IEPs.
	Impact on student learning		Teacher saw big difference and impact on student learning through HQIM.
	Student collaboration		Student collaboration with peers and teacher.
	Students working on disciplinary concepts		Students working on disciplinary concepts and skills articulated in the curriculum.
Student engagement		**5**	
	Entryways in HQIM		Number talks provided entryways for students.
	Student feedback on stories		Students seemed to really love the stories.
	Engaging students		Teachers were doing a better job of engaging students.
	Class discussion		Student-to-student and whole class discussions
	Multimedia		Multimedia provided entry points for all levels of learners.
Trainings		**3**	
	Training in math recovery		The teacher took training in math recovery before adopting HQIM.
	Professional development		Professional development on understanding each module
	Science of reading		Teacher was studying the science of reading.
Confidence in HQIM		**2**	
	Confident with HQIM		Teacher felt very confident with the curriculum. Teacher felt a step further in Year 3 with discourse.

	Still getting comfortable with HQIM		Teacher was still getting comfortable with some of HQIM and felt most comfortable in Year 4, when the teacher needed coaching support less and less.
HQIM design and implementation		**2**	
	Very pleased		Teacher was very pleased with the HQIM.
	Happy with rollout		Teacher was happy with rollout of HQIM and felt supported.
Leadership support		**2**	
	Coaching support		Coaching support was provided when requested by the teacher.
	Tiered implementation teams		Tiered implementation teams provided support for all teachers.
Resources		**2**	
	New manipulatives		Teacher felt that the new manipulatives received would be very effective.
	Materials for implementation		Support in having all materials needed for implementation
Building vocabulary		**2**	
	Math vocabulary		Math vocabulary was a daily part of lessons.
	Vocabulary development		Development of student vocabulary
Culturally responsive sustaining practices		**1**	
	Making connections to Dr. Martin Luther King, Jr.		Teacher decentered whiteness by making connection to Dr. Martin Luther King, Jr. during the lesson.
Data-informed decision-making		**2**	
	Exit tickets		Teacher used exit tickets to determine whether students met learning objective for the day.
	Assessment data		Assessment data were useful in making student groups for each unit.
Teacher preparation time		**2**	
	Ease of planning		Lesson planning was straightforward and not too time consuming.

	Additional preparation time		Teacher got to spend more becoming familiar with the lessons and doing more intellectual prep.

Note: "MCAS" stands for "Massachusetts Comprehensive Assessment System."

For the challenges reported with HQIM implementation, the most salient ones regarded *pacing instruction and the existing pacing guides* (9 out of the 18 observations*), lack of teacher collaborative time* (4 out of 18), and *challenges with the available resources for HQIM* (4 out of 18).

Table 29 provides details about the documented challenges, the frequency of occurrence in the sample of 18 teachers, and examples from the teacher interviews.

Table 29. Teacher reported challenges with HQIM implementation

Challenge Name	Description of Challenge	Frequency of Occurrence	Example from Teacher Interview
Pacing guide		**5**	
	Determining what materials to prioritize		It was a challenge to cut material out without losing the main idea and losing practice time.
	Cutting work leads to less practice		Cutting some of the work out led to less practice with math outside of assigned practice and problems.
	Some lessons not covered		A quarter of the curriculum did not get covered.
	Deciding which activities meet lesson goals		Lessons were adapted to fit time block. Activities were reviewed to see which met the learning goal.
	Time spent assessing		Assessing wasted a lot of learning time.
Pacing		**4**	
	Too fast		Pacing was fast, leaving less opportunities to practice skills.
	Feels rushed		Pacing felt rushed even after restructuring.
	Not much wiggle room to revisit material		Pacing was a challenge for students who were not where HQIM expected. At the beginning of the year, students struggled with solving equations. There was no time to go back and reteach.
	Pacing is a challenge		Sometimes lessons were shortened.
Teacher collaboration		**4**	
	Internalization of lesson plans takes long time		HQIM was a lot of work. Internalization process of lesson plans took a long time.

	Planning time spent in meetings		Teacher had 2 hours daily for planning. Some of the time was eaten up by meetings and planning with other teachers.
	Time for comfort in HQIM		Length of time to feel comfortable with HQIM.
	Getting used to new HQIM		It was challenging getting used to the new HGIM structure.
Resources		**4**	
	Not enough to fill time		Teacher had to find things to fill the time. Agile Minds lessons were 45 minutes, and a renaissance block was 65–70 minutes.
	Does not have all the materials to complete lesson		Consumable workbooks were not purchased. Teacher had to make copies of workbook pages daily.
	Students struggle with Chromebooks		Problems arose with students when using Chromebooks. Students struggled with using Chromebooks to use digital Illustrative Math tasks.
	Photocopying workbooks		Grant funding was running out. Teachers were photocopying workbooks.
Instructional content limits of HQIM		**3**	
	HQIM goes beyond standards		HQIM assessments went beyond standards.
	HQIM alignment with standards		HQIM was not aligned with grade-level standards.
	HQIM problem solving process		HQIM focused on problem-solving. There was no rote memorization in multiplication and division.
Engagement		**3**	
	Student connection to text		Students complained about some of the text and not connecting with the stories.
	Student Stamina		Lessons were often teacher directed. Students didn't have the stamina to sit as long.
	Connecting to vocabulary		Making connections to vocabulary for multilingual learners was a challenge.
PD/Trainings		**3**	
	Not much training on assessments		Assessments were long and not kid friendly. Teachers did not have much training on assessments.
	Not enough practice with genre writing		Genre writing was a struggle—not enough practice.

	PD support decrease		PD support decreased each year of implementation.
Culturally responsive sustaining practices		**1**	
	Challenge incorporating with current HQIM		It was hard to build in cultural responsiveness when it was not built into the curriculum.
Teacher turnover		**1**	
	3–4 different teachers in a school year		Students did not have consistency in 6th grade math with three to four different teachers.
Leadership support		**1**	
	Less support from leaders		There was less support from campus leaders with modifications to HQIM and co-planning.
Scaffolding		**1**	
	Scaffolding not in HQIM		Scaffolding was not in workbooks for multilingual learners.

Teachers reported a variety of supports that helped in the implementation of HQIM. The most frequently reported supports were around *training and professional development offered for HQIM* (12 out of18 observations), *coaching or instructional specialist support* (9 out of 18), *availability of coaches* (6 out of 18) and *having collaborative lesson planning time* (5 out of 18). Table 30 below provides details about the supports, the frequency of occurrence in the sample of 18 teachers, and examples from the teacher interviews.

Table 30. Teacher reported supports for HQIM implementation

Support name	Description	Frequency of Occurrence	Example from Teacher Interview
Training sessions/PDs	Curriculum-specific trainings, often prior to implementation or start of year	**12**	PDs for the last two years were held on half days and at the beginning of the year.
Regular meetings	Meetings with coaches, administrators or curriculum providers, sometimes as part of a coaching cycle	**9**	Teacher had consistent support from the math coach in addition to 40 minutes monthly support regarding questions, guidance, and/or lesson plan adjustments.
On-demand or reactive support	Administrators available to support, 24-hour support hotline	**6**	Coach was available when needed.

Collaborative lesson planning	Co-planning with coach or other teachers	**6**	Math coach co-planned with the teacher with the needs of all students in mind.
Materials	How to find and access materials online, development of supplemental materials	**5**	Instructional leadership specialist provided materials listed by Agile Minds to teachers when needed and informed teachers on how to find materials virtually.
Model lessons	Demonstrations of model lessons by coach or other teachers	**5**	Literacy coach modeled a lesson for teachers to observe.
Observations	Observation and feedback of their own teaching	**3**	In school observations and meetings regarding the program were held.
Assessments	Coaching or other support with revising or developing assessments	**3**	There were coaching supports in teasing out items beyond the standards and revising all assessments for better alignment with MCAS questions.
Co-teacher	Teacher and special education co-teacher	**1**	Having a co-teacher in special education was huge. Teacher and special education teacher taught entire days together. The special education SPED teacher embedded student accommodations into whole group instruction.

As part of teacher interviews, the Evaluation Team asked teachers about PD support they had received around culturally responsive teaching and learning. A majority of teachers (N = 6) reported having received one PD opportunity (e.g., workshop, session) during HQIM implementation. The PD opportunities were not necessarily tied to the HQIM and were more of an overview of what culturally responsive teaching and learning entail. Table 31 provides details about these PDs, the frequency of occurrence in the sample of 18 teachers, and examples from the teacher interviews.

Table 31. Teacher reported professional development opportunities around culturally responsive teaching and learning

Frequency of Receiving PD	Frequency	Example
One time	**6**	New teacher workshop is offered to discuss general information at [this school]. They go over being emotionally and culturally responsible.
Bimonthly	**2**	There was no PD specific to the curriculum. A lot of work around cultural responsiveness and bias was done in staff meetings and PDs as a school building.
Annually	**2**	ICs facilitated a PD on cultural sensitivity specifically for the HQIM curriculum.

Not PD, but integrated into curriculum or part of what administrator references	**2**	There was no PD at the district level. Specialists who conducted classroom observations regularly referenced culturally relevant teaching and equity as part of their observation.
Monthly	**1**	The teacher sat in on one module [that is part of a monthly series] and looked at culturally responsive teaching through that lens.
6 modules over school year	**1**	Three modules in a six modules program that were held throughout the year that included cultural responsiveness.
Frequent PDs offered but not attended	**1**	Frequent PDs were offered, but the teacher didn't sign up for them or remembered the specifics because they didn't have the capacity to attend.
Curriculum-specific PD around equity	**1**	Teacher participated in curriculum-specific PD around equity, including equity of using higher-level texts for all students and the rationale for that. Cultural responsiveness was discussed briefly, but it was not the focus of the PD.

Vignettes around Culturally and Linguistically Sustaining Practices

The EDC team drafted vignettes that showcased culturally and linguistically sustaining practices in the classrooms across MA DESE's categories: (1) asset-based teaching, (2) high expectations and support, (3) cultural competence and community building, and (4) partnerships with students and families. There were no vignettes under the category of sociopolitical awareness. Table 32 outlines example vignettes for each category.

Table 32. Vignettes of teachers' instructional practices around culturally and linguistically sustaining practices

Category	Category Definition	Frequency of Observation & Vignette Example
Asset-based teaching	Leveraging students' funds of knowledge (based on their cultures, lived experiences, and linguistic resources) as assets to support learning	Frequency: 6 Example: Connects learning to students' lives by discussing how nutrients give them energy, help them to run, play, think
High expectations and support	Supporting all students to develop positive identities as learners, attain the academic skills and knowledge to meet or exceed grade-level standards, and apply competencies in relevant, real-world contexts	Frequency: 18 Example: Uses affirming language while prompting students to explain their thinking Example: Supports the learning of all students using an ASL interpreter and a translation app for students who can read in Spanish
Cultural competence and community building	Creating a learning environment that is affirming of diversity and where each student feels a sense of belonging while developing respect and understanding for cultures and	Frequency: 3 (all were emerging practices, not fully developed) Example: Drawing parallels in history by connecting lesson topic (Bessie Coleman) to other culturally relevant historical figures and themes like MLK, Jr. and segregation

	identities that are different from their own	
Sociopolitical awareness	Empowering students with the ability to identify, analyze, and work to solve real-world problems by thinking critically and drawing conclusions about and developing agency around complex issues, including those related to equity, identity, power or bias	Frequency: 0 Example: No vignettes
Partnerships with students and families	Incorporating student/family voice and creating opportunities for meaningful engagement in the classroom community and learning process	Frequency: 14 Example: Students use each other as resources in turn and talk. Teacher encourages students to show multiple ways of solving a problem

Document Review: Landscape Analysis

Each district, with the help of their IC, conducted a landscape analysis of their current implementation of the HQIM as one of the initial grant activities. These landscape analyses assessed systems and structures in place to support implementation as well as teacher FOI and leadership support.

The Evaluation Team reviewed each district's landscape analysis to provide feedback to districts on the strengths of their landscape analysis and on the areas that were not examined in their landscape analysis that might help with their implementation.

The team's landscape analysis review process involved applying factors identified in a review of the literature to develop key elements to look for in each landscape analysis. As part of this process, the team also examined each IC's framework for conducting landscape analyses to ensure alignment with the Evaluation Team framework. (See the Landscape Analysis Review Checklist in Instrument 9.)

Analysis Procedures

The Evaluation Team reviewed the landscape analysis for each school using a checklist aligned to the framework of HQIM implementation. Each landscape analysis was reviewed by at least two Evaluation Team researchers. MA DESE was provided with a summary of the feedback about which elements were evident—or not evident—in the landscape analyses. Table 33 provides a summary of the results of the reviews.

Table 33. Landscape analysis review results

Elements	Findings
Vision and selection	21 districts (40%) had a content-specific instructional vision.

	• 20 districts (38%) used a collaborative process for HQIM selection.
Curriculum-embedded professional learning	• 34 districts (65%) had a curriculum-embedded professional learning: ○ 24 districts (46%) had effective coaching practices. ○ 20 districts (38%) effectively used professional learning community time.
Leadership support	• 29 districts (56%) identified leadership supports for implementation: ○ 12 districts (23%) communicated clear expectations for HQIM implementation. ○ 9 districts (17%) had dedicated resources for professional learning. ○ 6 districts (12%) conducted admin walkthroughs using a curriculum or vision-aligned tool. ○ 5 districts (10%) communicated consistently about HQIM implementation. ○ 3 districts (6%) provided resources for continuous improvement support.
Systems and structures	• 17 districts (33%) had systems and structures for continuous improvement. • 15 districts (29%) had small group/intervention structures utilizing aligned materials. • 10 districts (19%) had sufficient time in the schedule allocated for the HQIM.
Racial equity	• 4 districts (8%) offered culturally-responsive instruction or anti-bias professional learning for staff. • 3 districts (6%) examined disaggregated data for improvement.
Teacher buy-in	• Teachers in seven districts (13%) have begun to believe in the HQIM as a result of early student successes with the curriculum.

Document Review: Implementation Plans

Each district, with the support of their IC, developed an implementation plan based on the results of their landscape analysis. These plans identified priority goals for the year and action steps for achieving these goals.

The Evaluation Team reviewed each district's implementation plan to provide feedback to districts on strengths of their plans and on the areas that were not examined in their plans that might help with their implementation.

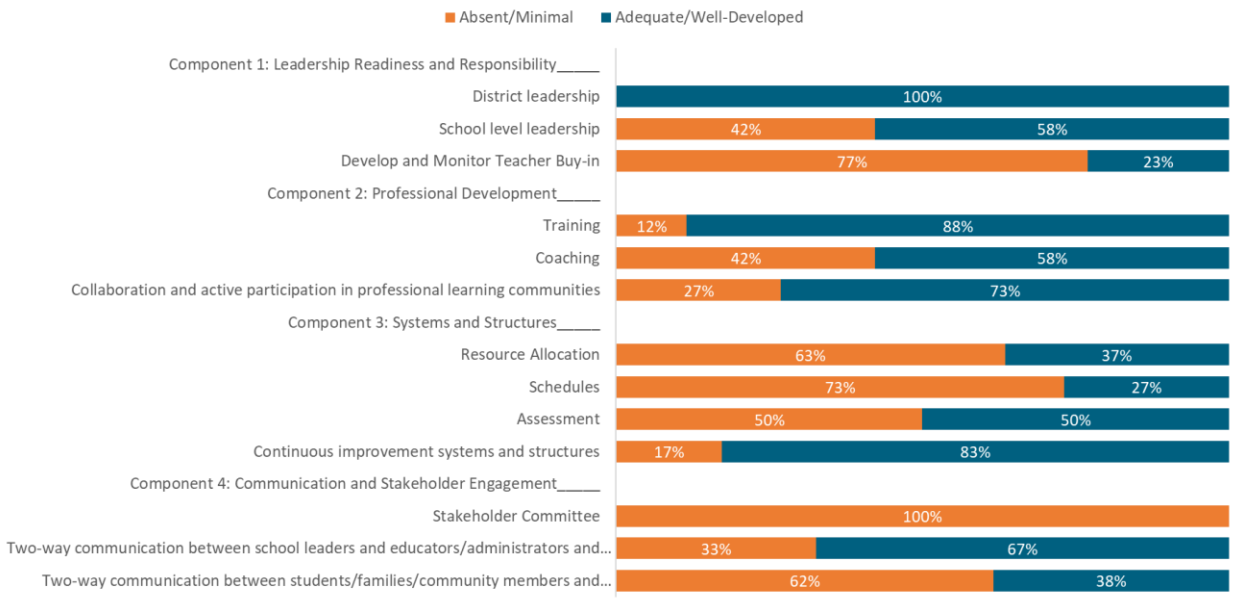
Analysis Procedures

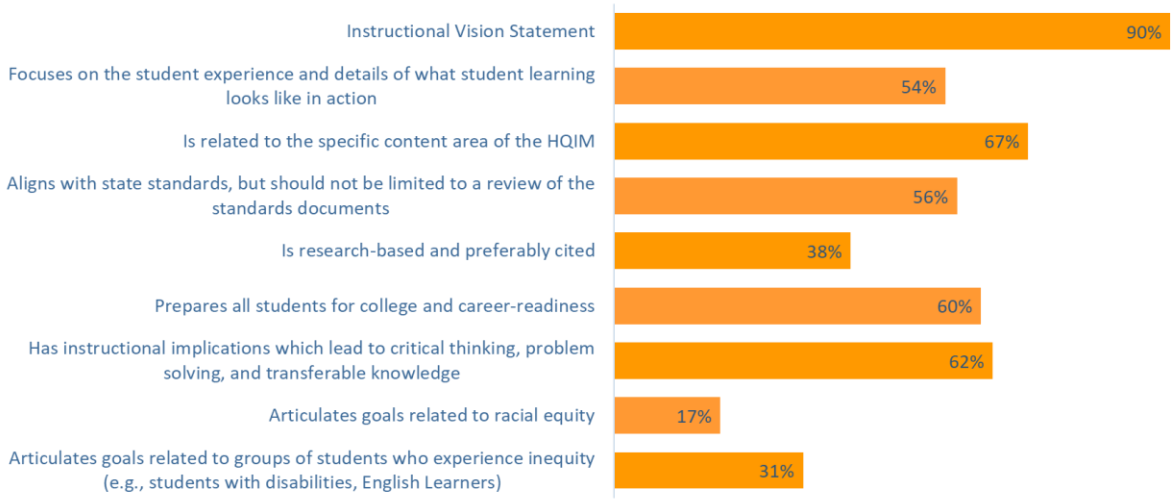
The Evaluation Team created a Quality of Implementation Plan Rubric (See Instrument 10. Quality of Implementation Plan Rubric) that listed MA DESE's requirements as well as four components that were essential to the implementation of HQIM (as based on research). Each component comprised a number of elements that were categories of observable, testable items. Listed under each element were action steps that were the evidence-based "look-fors" that a well-developed plan might include.

For each implementation plan, a pair of Evaluation Team researchers independently used the rubric to indicate which components and elements were being addressed in the plan, with particular attention to how the plan centered racial equity and building leadership capacity. Each element was rated as "not addressed," "somewhat addressed." or "fully addressed." The Evaluation Team researchers then met to calibrate their ratings and consolidate the comments. The completed rubric was then reviewed by the project lead and any needed revisions were made.

All of the ratings were displayed in a scorecard along with a presentation of exemplary implementation plan components, both of which were shared with MA DESE. See Figure 2 to view a summary of the results of the reviews of the grantee Implementation Plans.

Figure 2. Graph of results of reviews of grantee Implementation Plans




Instructional Vision Statement
90%
Focuses on the student experience and details of what student learning looks like in action
54%
Is related to the specific content area of the HQIM
67%
Aligns with state standards, but should not be limited to a review of the standards documents
56%
Is research-based and preferably cited
38%
Prepares all students for college and career-readiness
60%
Has instructional implications which lead to critical thinking, problem solving, and transferable knowledge
62%
Articulates goals related to racial equity
17%
Articulates goals related to groups of students who experience inequity (e.g., students with disabilities, English Learners)
31%

Instruments

The Evaluation Team has compiled a set of 12 instruments and other documents developed during the project into a separate Zip file, which accompanies this report. This set includes final versions of each survey, interview and observation protocol, and materials used for document review. Some of these materials are referred to in the technical report. The full list of instruments is provided here as reference.

1. Teacher survey (Time 2)
2. DIGL survey
3. IC perspectives survey
4. IC interview protocol- Time 1
5. IC interview protocol- Time 2
6. Teacher observation protocol
7. Teacher interview template
8. FOI matrix rating sheet
9. Landscape analysis rubric
10. Implementation plan rubric
11. Qualitative codebook

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Charmaz, K. (2014). Grounded theory in global perspective: Reviews by international researchers. *Qualitative Inquiry, 20*(9), 1074–1084.

Gay, G. (2000). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Grant, C. A., & Gillette, M. (2006). A candid talk to teacher educators about effectively preparing teachers who can teach everyone's children. *Journal of Teacher Education, 57*(3), 292–299.

Hackman, H. W. (2005). Five essential components for social justice education. *Equity & Excellence in Education, 38*(2), 103–109.

Kane, T. J., Owens, A. M., Marinell, W. H., Thal, D. R. C., & Staiger, D. O. (n.d.). *Teaching higher: Educators' perspectives on common core implementation*. Center for Education Policy Research, Harvard University. https://cepr.harvard.edu/files/cepr/files/teaching-higher-report.pdf

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491.

Martin, M. O., & Mullis, I. V. (2013). *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—Implications for early learning*. International Association for the Evaluation of Educational Achievement. https://omb.report/icr/201011-1850-001/doc/21114001.pdf

Massachusetts Department of Education. (2024). *A resource for culturally and linguistically sustaining practices*. Internal document reproduced with permission from MA DESE.

Mullis, I. V., Martin, M. O., Fishbein, B., Foy, P., & Moncaleano, S. (2021). *Findings from the TIMSS 2019 problem solving and inquiry tasks*. Boston College, TIMSS & PIRLS International Study Center. https://nces.ed.gov/timss/pdf/T19_Gr8_MthTchQ_USA_Questionnaire.pdf

National Center for Education Statistics. (2021). *National teacher and principal survey*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/surveys/ntps/pdf/2021/Teacher_Questionnaire_2020_21.pdf

Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). Exploratory evaluation. *Handbook of practical program evaluation* (4th ed, pp.88–107). Jossey-Bass.

Rodriguez, J. L., Jones, E. B., Pang, V. O., & Park, C. D. (2004). Promoting academic achievement and identity development among diverse high school students. *The High School Journal, 87*(3), 44–53.

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). Sage Publications.

U.S. Department of Education, Institute of Education Sciences. (2019). *TALIS 2018 U.S. highlights web report* (NCES 2019-132 and NCES 2020-069). https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019132

Waxman, H. C., & Tellez, K. (2002). *Research synthesis on effective teaching practices for English language learners*. Publication Series.

Appendices

The Evaluation Team has compiled the results from the multilevel models run for MRQ3 and LRQ1 into two appendices. These appendices along with other documents developed during the project into a separate Zip file, which accompanies this report.

- Appendix A. Teacher Survey Time Two Multi-level Models for MRQ3
- Appendix B. Hierarchical Linear Modeling Analysis (HLM) of LRQ 1