Cohort Wide Learning Plan

July 2021

Prepared for the Bill & Melinda Gates Foundation by
Caryn S. Ward, Ximena Franco-Jenkins, and Collin McColskey-Leary
# TABLE OF CONTENTS

## CONTENT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>3</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>9</td>
</tr>
<tr>
<td>RESEARCH DESIGN</td>
<td>10</td>
</tr>
<tr>
<td>OUTPUTS</td>
<td>27</td>
</tr>
<tr>
<td>METHODS</td>
<td>32</td>
</tr>
<tr>
<td>SAMPLE</td>
<td>32</td>
</tr>
<tr>
<td>MEASURES</td>
<td>33</td>
</tr>
<tr>
<td>PROCEDURES</td>
<td>45</td>
</tr>
<tr>
<td>DATA ANALYSIS PLAN</td>
<td>48</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>54</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>58</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>61</td>
</tr>
</tbody>
</table>
**EXECUTIVE SUMMARY**

**PURPOSE**

The overall aim of the Effective Implementation Cohort (EIC) investment is to increase district capacity to implement a high-quality middle years math curriculum as part of a Coherent Instructional System (CIS) to accelerate learning for students who are experiencing poverty, Black, Latino/a, and/or English Learner (EL)-Designated (“priority students”). As a Learning Partner, NIRN seeks to support partnerships between Providers and Local Education Agencies (LEAs) in their implementation and measurement efforts, as well as to collect and study data to answer the investment’s learning questions within the cohort wide learning agenda. Implementation support activities and data collection, analysis, and reporting activities for the cohort wide learning agenda will be guided by the Cohort’s learning questions and theory of action grounded in implementation science.

**THE BILL AND MELINDA GATES FOUNDATION’S LEARNING QUESTIONS ARE GROUPED INTO FOUR CATEGORIES:**

- **Application of Established Practice**
- **Variability in Implementation**
- **Enabling Contexts**
- **Measurement**

**DESIGN**

Our research design examines how degrees of and factors related to implementation influence a teacher’s ability and beliefs to implement a curriculum as intended (fidelity), and ultimately, how teacher self-efficacy and the level of fidelity of implementation affect students’ mathematics engagement, experience, beliefs, and achievement. A mixed-methods approach will be used to measure the strength and variation of implementation supports across Provider-LEA partnerships (n = 19) at the district, school, and classroom/teacher levels in order to identify cohort-wide trends and establish relationships between implementation outcomes and other contextual factors (i.e., locale, governance model, and leadership stability).
Various implementation supports are provided by NIRN, Providers, and LEAs. Multiple methods of data collection will be used to gather information from LEA and school leaders, teams (inclusive of district and school staff), math teachers, instructional coaches, students, and providers. Specifically, data collection methods will include teacher and student surveys, observations, interviews, administrative data, and product reviews. Common data collection methods across the cohort will be used to examine students’ beliefs, experience, and engagement in mathematics (i.e., student survey) and the LEA’s and school’s organizational capacity (e.g., District Capacity Assessment, Principal Implementation Leadership Survey). For other implementation outcomes such as teacher knowledge, self-efficacy, and practice, Providers and LEAs will be using instruments that best fit their individual work and context.

Thus, the analysis will be largely descriptive using cohort-wide trends to achieve goals including:

- Creating evidenced-based models for the implementation of high-quality middle years math curricula
- Adding to the knowledge base available to education professions on how to implement high-quality middle years math curricula, particularly for students of color and those experiencing poverty
- Providing participating Providers and LEAs with meaningful, formative feedback to inform their practice
DEEP DIVE IMPLEMENTATION STUDY

To further investigate the learning question of how does implementation affect student math learning and for whom and in what contexts, student and teacher level unit of analysis will be conducted with a select number of Provider-LEA partnerships (i.e., up to 4 to 6 LEAs). The sub-sample may include LEAs also participating in the AMS study of efficacy and enactment, depending on the final district/school selection and the student and teacher data being collected as part of that study, as well as those LEAs with the ability to provide linked student and teacher administrative data. Efforts will be taken to ensure the sub-sample of LEAs is representative of key contextual factors such as representation of priority students and LEA locale. Quantitative analyses for this sub-sample will utilize propensity score analysis and multilevel linear modelling. Although limitations to generalizability are present with this deep dive approach, controlling for school and teacher factors that might influence implementation fidelity and student outcomes will allow the isolation of the effect of implementation on student learning, for whom, and in what contexts. Given the purpose of this investment is an implementation learning agenda and does not seek to establish causal relationships, a control group will not be used.
OUTPUTS

In addition to improving LEA and Provider capacity, the EIC has the goal of producing reliable, practical evidence and measures to inform planning and implementing a district-wide, high-impact math improvement initiative. The evidence and measures will support:

- **Providers to design more effective implementation services**
- **Districts as they track implementation progress**
- **The field to better understand how high-quality implementation affects enactment, and, in turn, student learning**

A number of potential outputs will be generated from the EIC for use by the field of education and in other BMGF investments related to the implementation of middle years mathematics curriculum as part of a Coherent Instructional System.
Specifically, the design will use the school level as the primary unit of analysis to examine all of the learning questions. Qualitative and quantitative data will be conducted through descriptive, visual trends, and thematic analyses. **Example outputs of the cohort wide learning agenda using this design include:**

<table>
<thead>
<tr>
<th>Definition of readiness conditions at school and district levels that support scaling effective solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A curated list of self-assessment tools for district leaders to gauge their system’s readiness and identify next steps to prepare for implementation</td>
</tr>
<tr>
<td>Identification of context-specific readiness factors that leaders in priority contexts should consider</td>
</tr>
</tbody>
</table>

Example outputs of the cohort wide learning agenda using this design include:
INTRODUCTION

The overall aim of the EIC investment is to increase district capacity to implement high-quality middle years math curricula as part of a Coherent Instructional System to accelerate learning for students who are experiencing poverty, Black, Latino/a, and/or English Learner (EL)-Designated (“priority students”). As a Learning Partner, NIRN seeks to support partnerships between Providers and Local Education Agencies (LEAs) in their implementation and measurement efforts, as well as to collect and study data to answer the investment’s learning questions within the cohort wide learning agenda.

THE BILL AND MELINDA GATES FOUNDATION’S LEARNING QUESTIONS ARE GROUPED INTO FOUR CATEGORIES:

- APPLICATION OF ESTABLISHED PRACTICE
- ENABLING CONTEXTS
- VARIABILITY IN IMPLEMENTATION
- MEASUREMENT

It is important to note that the EIC investment consists of two phases.

**Phase I (January-July 2021)**
Was focused on implementation planning and creating readiness for high-quality middle years mathematics curriculum.

**Phase II (August 2021-June 2024)**
Is focused on implementation of the high-quality middle years math curriculum.
RESEARCH DESIGN
Learning Questions

This investment aims to answer questions on the application of established practice and enabling contexts beginning in Phase I and continuing to answer those questions in Phase II. The learning questions related to variability in implementation and measurement will be answered during Phase II.

The specific questions under each category are detailed below:

**APPLICATION OF ESTABLISHED PRACTICE**

1. Which lessons learned from implementation research (e.g., curriculum enactment studies, education and health implementation studies) scale across this set of partnerships? Which do not?
2. What do educators and system actors identify as the reasons why evidence-based ideas scale or do not scale?

**ENABLING CONTEXTS**

1. What conditions most distinguish districts’ and schools’ readiness for, approach to, and success with implementation?
2. What conditions maximize likelihood of successful implementation across contexts relevant to priority students?
3. Which conditions are go/no-go factors versus important readiness conditions that a provider can help establish?
4. How does the set of key enabling conditions vary across priority contexts?
5. How have districts’ priorities for implementation of new curricula or instructional systems shifted in light of COVID-19? How do these shifts influence key enabling conditions?
1. What is the “menu” of interventions taken by systems to support implementation? Are some seen as more promising and likely to be taken up than others? Which seem to achieve successful implementation most often? How does this vary by context?

2. What are the challenges for implementation faced across grantees? What patterns are there to those challenges, especially as related to contextual factors (e.g., student composition, characteristics of previous instructional system, teacher characteristics, broader enabling conditions)?

3. What best practices emerge across the dimensions of implementation support that lead partners are being asked to provide, as well as any others that arise? How are those best practices related to patterns in contextual factors as described above?

4. How do the findings from this cohort reinforce, complement, or differ from the findings about classroom-level factors that drive curricula’s effectiveness in the separate enactment studies that the Foundation is funding?

1. What defines effective implementation? What are evidence-based leading indicators, lagging indicators, and data sources to assess progress and impact?

2. What are the best-established measures for assessing readiness to take on implementation, self-assessment of what is needed to support implementation, and factors (e.g., system conditions) that should be tracked to support implementation?

3. How do grantees use or eschew established leading and lagging measures related to both CIS and implementation to identify challenges and make improvements to systems of support? Why? We are particularly interested in findings related to the ability of systems to collect and share data, perceptions of the validity of measures, and actionability of data.

4. What factors not identified in established measures are critical to identifying challenges and making improvements? How do districts and school systems identify these factors?

5. How, and how much, did implementation affect student math learning and/or leading indicators of math learning? In what contexts and for whom?

6. How does teacher perception (e.g., feasibility, acceptability, and appropriateness) of the math curriculum within CIS influence implementation?
RESEARCH DESIGN
Implementation Supports

Our research design examines how degrees of and factors related to implementation influence a teacher’s ability and beliefs to implement the curriculum as intended (fidelity), and ultimately, how teacher self-efficacy and the level of fidelity of implementation affect students’ mathematics engagement, experience, beliefs, and achievement. A mixed-methods approach will be used to measure the strength and variation of implementation supports across Provider-LEA partnerships (n = 19) at the district, school, and classroom/teacher levels in order to identify cohort-wide trends and establish relationships between implementation outcomes and other contextual factors (i.e., locale, governance model, and leadership stability).

Specifically, the research design examines interconnected implementation supports operating at different levels of the system. Each layer of implementation support is designed to increase capacity to implement effective middle years math curriculum, with the ultimate goal of increasing student achievement particularly for students of color and those experiencing poverty. Implementation supports vary based on Provider and LEA needs.

The three levels of implementation supports are:

1. NIRN’S SUPPORT TO PROVIDERS AND LEAS AS THE LEARNING PARTNER
2. PROVIDERS SUPPORT FOR LOCAL EDUCATION AGENCIES
3. PROVIDER/LEA SUPPORT FOR TEACHERS AND SCHOOL LEADERS
A cohort wide theory of action will be used to examine the cohort’s learning questions. Specifically, the theory of action outlines how change will lead to improved practices and outcomes and guides the development of hypothetical statements. The proposed theory of action is based on implementation science that posits three factors are necessary for improved outcomes including: 1) effective practices (e.g., high-quality middle years math curriculum and instructional practices), 2) effective implementation (e.g., high quality professional learning services), and 3) enabling context (e.g., linked implementation teams and use of improvement cycles) (Fixsen et al., 2013). See Figure 1 for the Theory of Action.

Using the theory of action, a model outlining the enabling conditions (i.e., readiness indicators) and series of short-term, intermediate, and long-term outcomes has been developed to inform the learning questions (see Figure 2 p.14). In addition to the readiness indicators, a number of district context factors have been identified to be examined.
### FIGURE 2.
Map of Theory of Action & Outcomes

<table>
<thead>
<tr>
<th>PHASE 1</th>
<th>PHASE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LQ1</strong> Application of Established Practice</td>
<td><strong>LQ1</strong> Application of Established Practice</td>
</tr>
<tr>
<td><strong>LQ2</strong> Enabling context</td>
<td><strong>LQ2</strong> Enabling context</td>
</tr>
<tr>
<td><strong>LQ3</strong> Variability in Implementation</td>
<td><strong>LQ3</strong> Variability in Implementation</td>
</tr>
<tr>
<td><strong>LQ4</strong> Measurement</td>
<td><strong>LQ4</strong> Measurement</td>
</tr>
</tbody>
</table>

#### IF
**Enabling Conditions**

#### THEN

**Professional Learning Services (PLS) & High-Quality Math Curriculum**

<table>
<thead>
<tr>
<th>Short Term</th>
<th>Intermediate</th>
<th>Long Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>Cost</td>
<td>Student Math Achievement (benchmark, formative and summative assessments)</td>
</tr>
<tr>
<td>Adoption</td>
<td>Appropriateness</td>
<td>Student Beliefs (enjoyment, self-efficacy, and growth mindset)</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>Feasibility</td>
<td>Teacher Self-Efficacy: Math Curriculum Cultural Pedagogy</td>
</tr>
<tr>
<td>Feasibility</td>
<td>Quality of PLS</td>
<td>Teacher Practice: Fidelity/ Integrity</td>
</tr>
<tr>
<td>Quality of PLS</td>
<td><strong>Intermediate</strong></td>
<td><strong>Long Term</strong></td>
</tr>
<tr>
<td><strong>Intermediate</strong></td>
<td>Teacher Practice: Fidelity/ Integrity</td>
<td><strong>Long Term</strong></td>
</tr>
<tr>
<td>Intermediate</td>
<td>Teacher Practice: Fidelity/ Integrity</td>
<td>Long Term</td>
</tr>
<tr>
<td><strong>Intermediate</strong></td>
<td>Teacher Practice: Fidelity/ Integrity</td>
<td>Long Term</td>
</tr>
</tbody>
</table>

#### District Contextual Factors:
- **Locale**
- **Size**
- Demographics of Student Population
- Governance Structure (i.e., elected vs appointed, relation with school board)
- LEA Leadership Stability (i.e., Superintendent, Chief Academic Officer)
- Decision Making Model (i.e., centralized or decentralized - site based)
- School Leadership Stability
- Teacher Retention
- Labor Relations (leadership & union agreement on strategy)
- Community Relations (engagement of school board, engagement of regional/state support)
- Financial Management (i.e., funding model - braided, transferred, blended)
SUMMARY OF EMPIRICAL EVIDENCE

A summary of the existing empirical evidence for different theory of action constructs is provided. In addition to highlighting the evidence of what is known, we also highlight how the proposed study will contribute to or generalize new knowledge regarding these constructs for the field.

ENABLING CONDITIONS

IF

If establish linked teams to create enabling conditions using the sciences of implementation improvement at the Provider, district, and school levels

Our theory of action begins with the premise that the foundation for improved outcomes for all students in mathematics is the establishment of linked teams and the creation of enabling conditions at district and school levels to support the use of a high quality middle years mathematics curriculum within a coherent instructional system. Using the sciences of implementation and improvement, we will examine the role and functioning of a linked teaming structure (provider - district - school - level teaming) to create and maintain the necessary readiness conditions for implementation.

Readiness is defined as a developmental point at which a person, organization, or system has the capacity and willingness (momentum) to engage in a particular activity (Fixsen et al., 2013; Wandersman & Scaccia, 2018).
“Willingness” is defined as the quality or state of being prepared to do something.

“Capacity” is defined as the perceived abilities, skills, and expertise of school leaders, teachers, faculties to execute or accomplish something specific, such as leading a school-improvement effort.
Readiness is a condition that needs to be developed and maintained throughout implementation. It is often considered to be a precursor to successful implementation but has not been as widely researched as individual readiness for change (Weiner, 2020). Social cognitive theory and motivational theory supports the following hypotheses underlying our theory of action: *When organizational readiness for change is high, organizational members are more likely to initiate change, exert greater effort to implement the change, and persist in change implementation despite obstacles* (Gist & Mitchell, 1992; Meyer & Herscovitch, 2001).

Specifically, Provider-LEA pre-implementation readiness conditions will be examined across six constructs:

- **District Teaming (DT)**
- **LEA Executive Sponsor Engagement (ES)**
- **Communication (C)**
- **Assessing Fit and Feasibility (FF)**
- **Implementational Planning (IP)**
- **Measurement Planning (MP)**

Each readiness construct consists of 4 to 6 operationalized indicators contextualized for this specific investment (See Appendix B for all readiness construct definitions and indicators). A summary of the evidence base for each of these readiness constructs is provided in Table 1 (p.18).
### TABLE 1.
Readiness Constructs’ Definitions & Rationales

<table>
<thead>
<tr>
<th>Readiness Construct</th>
<th>Definitions</th>
<th>Rationale</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>District Teaming</strong></td>
<td>A representative team has been formed at the district level and is working to lead implementation and create the enabling conditions in selected schools so that educators can make full use of high-quality middle years math curriculum.</td>
<td>Research has shown that using implementation teams to actively and intentionally make changes produces higher rates of success more quickly than traditional methods of implementation with less active approaches (Higgins et al., 2012; Metz et al., 2015). Research has also shown implementation teams have a significant impact on whether evidence based practices are implemented and sustained over time (Leithwood &amp; Azah, 2017; McIntosh et al., 2018). For mathematical instructional systems specifically, evidence exists to support the positive impact of school based professional learning communities and teacher networks on teacher learning (Cobb &amp; Jackson, 2011).</td>
<td></td>
</tr>
<tr>
<td><strong>LEA Executive Sponsor Engagement</strong></td>
<td>The Executive Sponsor(s) champions and supports district and school staff as they engage in implementation of high-quality middle years mathematics curriculum. An executive sponsor is an individual with the ability to influence others, authority to make decisions regarding resource allocation, institutional knowledge, time, and positive relationships with staff and stakeholders.</td>
<td>Executive sponsors or champions are defined as those individuals who are internal to the organization, have an intrinsic interest in the needed change, are driven to succeed, have a positive approach to the work, and a strong conviction or belief in their ability to do the work. Research has found that champions represent a “necessary but not sufficient” condition for implementation success. Champions alone are inadequate to bring about change, yet in combination with other factors were found essential to implementation success (Miech et al., 2018). The role of a “champion” or executive sponsor at the LEA level has not been researched specifically for mathematics. Instructional leadership provided by principals has been found in research to be critical for successful implementation (Cobb &amp; Jackson, 2011; Katterfeld, 2013).</td>
<td></td>
</tr>
</tbody>
</table>
Readiness Construct

Communications

| Definitions | Frequent and accurate information regarding implementation planning and progress is exchanged between stakeholders and acted upon by the identified persons. Goals of communication include sharing information, gathering feedback and input, clarifying expectations, and celebrating successes. |
| Rationale | Research has demonstrated that strategic communication to be an important driver for system change leading to improved performance (Fullan, 2010; Levine, 2014; Duffy & Chance, 2006). |

Readiness Construct

Assessing Fit and Feasibility

| Definitions | An assessment conducted by districts to better understand how a new or existing high-quality middle years mathematics curriculum works within their existing context to support implementation planning and use. |
| Rationale | Based on social cognitive theory, change efficacy is largely a function of organizational members’ cognitive appraisal of task demands, resource availability, and situational factors (Gist & Mitchell, 1992). “When organizational members share a common, favourable assessment of task demands, resource availability, and situational factors, they share a sense of confidence that collectively they can implement organizational change” (Weiner, 2020, p. 222). Although the importance of contextual variables and fit is often referenced in selection and implementation of evidence based practices, there is lack of consensus on specific elements that constitute contextual fit and lack of a strong research base (Horner, Blitz, & Ross, 2014). Research has shown that districts consider most prominently the three factors of alignment to standards, impact on student achievement, and cost effectiveness when selecting curriculum materials (Allen & Seaman, 2017; Campbell & Polikoff, 2017; Zeringue et al., 2010). |

Readiness Construct

Implementation Planning

| Definitions | Comprehensive implementation strategies are specified within a plan to ensure capacity is developed to support successful use and sustainability of a high-quality middle years mathematics curriculum. |
| Rationale | Powell and colleagues (2015) found having a formal implementation “blueprint” to guide implementation was a key implementation strategy for change. The plan or blueprint outlines the 1) aim/purpose of the implementation, 2) scope of the change, 3) timeline for change, and 4) appropriate performance/progress measures. Within K-12 education, quality school planning has been found to be associated with positive implementation outcomes (Strunk et al., 2016). |
**Readiness Construct**

**Measurement Planning**

<table>
<thead>
<tr>
<th>Definitions</th>
<th>Utilize multiple methods to collect and review data to: (1) inform decision-making for continuous improvement, (2) examine effectiveness, and (3) communicate with stakeholders.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rationale</td>
<td>Measurement planning and data usage have been found to be a district practice associated with successful implementation of evidence based practices (Byrk et al., 2015; Fullan &amp; Quinn, 2016; George et al., 2018; Leithwood &amp; Azah, 2017).</td>
</tr>
</tbody>
</table>

In addition to the identified readiness constructs, a number of district and school demographics and contextual factors (such as locale, governance structure, labor relations, community relations, and financial management) will be examined. Student and teacher demographic and socio-economic status has been found to be associated with student achievement (LaCour & Tissington, 2011; McCoy, 2005).

Furthermore, as the percentage of students living below poverty and the diversity among students increases, teachers’ perceptions of shared norms for instruction, climate, openness with parents, and district support decreases. Although larger urban districts have been found to foster greater shared leadership among teachers and staff, perceptions of school climate, openness with parents, and district support decreases as compared to teachers’ perceptions within suburban and rural districts (Louis et al., 2010).

In terms of students having a teacher of the same race, Redding (2019) found that students’ perceptions of a same-race teacher is more favorable. This relationship varies by the school level. There was also strong evidence that Black students perform higher on academic achievement tests when assigned to a Black teacher. There was less evidence of this relationship for Latino/a students.
PROFESSIONAL LEARNING AND MATH CURRICULUM IMPLEMENTATION OUTCOMES

IF/THEN Districts and schools will have increased capacity to select and implement high-quality middle years mathematics curriculum and professional learning supports. While significant financial and human resources are allocated to professional learning activities in districts, less is allocated to support the foundational components that facilitate implementation and scale-up. As displayed in Table 2, Proctor and colleagues (2011) identified a set of outcomes to understand what and how processes work to facilitate or create barriers to implementation. While conceptually sound, they have not undergone rigorous examination in education. The EIC body of work adds significantly to how districts and schools can better attend to, collect, analyze, and use implementation outcomes to increase the likelihood that teachers implement practices as intended.

TABLE 2.
Implementation Outcome Constructs’ Definitions & Rationales

<table>
<thead>
<tr>
<th>Implementation Outcome Construct</th>
<th>Definitions</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability and Appropriateness</td>
<td>Acceptability is defined as the perception that a given practice (i.e., math curriculum) or service (i.e., professional learning services) is agreeable and satisfactory.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Appropriateness is defined as the perceived fit, relevance or compatibility of a practice (i.e., math curriculum) or service (i.e., professional learning services) to address a specific need.</td>
</tr>
<tr>
<td></td>
<td>Rationale</td>
<td>While little research has established their predictive capability as implementation outcomes, Weiner and colleagues (2017) developed measures of the acceptability and appropriateness of interventions and practices in local contexts. Their work identified acceptability of an intervention or practice as being approved, appealing to, liked, and welcomed by practitioners. Furthermore, appropriateness was operationalized as being fitting, suitable, applicable, and a good match to the practitioners current work and environment.</td>
</tr>
<tr>
<td></td>
<td>Feasibility</td>
<td>Feasibility is the extent to which a practice (i.e., math curriculum) or service (i.e., professional learning services) can be successfully delivered or used within a given context.</td>
</tr>
<tr>
<td></td>
<td>Rationale</td>
<td>Lyon and others (2019) found that beliefs on the importance and feasibility of implementation of a practice can shift over time and should be addressed in implementation planning.</td>
</tr>
</tbody>
</table>
### Implementation Outcome Construct

#### Quality of Professional Learning

**Definitions**
The extent to which a service is perceived to be of high quality and usefulness.

**Rationale**
Quality of professional learning had the potential to impact both teacher practice and student outcomes. Brock and Carter (2016) found that modeling and performance feedback during training had a significant positive impact on fidelity, especially when it came to implementing programs for students with disabilities. Additional evidence has shown that job embedded training increases teacher-self-efficacy and student math achievement (Althauser, 2015).

### Implementation Outcome Construct

#### Organizational Capacity: Communication, Data Systems, Leadership Quality, Resource Allocation, Stakeholder Engagement, Supporting Policies & Procedures

**Definitions**
The systems, activities, and resources that are necessary for schools to successfully adopt and sustain practices (i.e., math curriculum) such as Communication, Data System, Leadership Quality, Resource Allocation, Stakeholder Engagement, and Supporting Policies & Procedures.

**Rationale**
Organizational capacity is critical to understanding the resources, processes, and structures that support teachers in implementing new practices and programs. Organizational activities such as communication, leadership, stakeholder engagement, and supportive policies and procedures that are aligned to the literature that suggests schools can enhance teacher’s involvement in professional learning and connection to the organization (Sleeger et al., 2014). Additionally, Malen and colleagues (2015) suggest “building capacity requires the timely allocation of appropriate levels and types of foundational resources and the strategic construction of the organizational conditions that allow those resources to be put to productive use.”

### Implementation Outcome Construct

#### Cost/Sustainability

**Definitions**
Cost of implementation efforts for the practice (i.e., math curriculum) or service (i.e., professional learning services)

Sustainability is the extent to which a practice (i.e., math curriculum) or service (i.e., professional learning services) is maintained within a system.

**Rationale**
Cost as an implementation outcome is important to both sustainability and replicability. Levin and Belfield (2015) suggest that evaluations should include a component that collects and analyzes cost information using the ingredients method, which specifies the activities that are needed to fully enact the program or practice. Estimate cost figures allow researcher to establish cost-effectiveness and cost-benefit estimates that can inform decision making.
THEN Teachers will have improved self-efficacy and integrity/fidelity of implementation of high-quality middle years mathematics curriculum.

The EIC theory of action posits with improved organizational capacity and enabling conditions in place, teachers will experience positive changes in their knowledge and skills in the delivery of mathematics instruction using the high quality curriculum materials as well as in their confidence in delivering instruction and meeting student needs including the use of culturally pedagogy strategies. Overall, research has shown that implementation strategies such as professional development and coaching can have an impact on teacher beliefs and practice (Desimone & Garet, 2015; Kraft et al., 2017). Furthermore, teachers’ beliefs about instruction and student learning influences their instructional practice as well as their response to change initiatives (Spillane, Hopkins & Sweet, 2017; Stipek, Givvin, Salmon, & MacGyvers, 2001). Teacher outcomes are displayed in Table 3.

**TABLE 3.**
Teacher Outcome Constructs’ Definitions & Rationales

<table>
<thead>
<tr>
<th>Teacher Outcome Construct</th>
<th><strong>Teacher Knowledge &amp; Skills of Math Instruction</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>The degree of knowledge and skills in teaching mathematics (delivery of instruction).</td>
</tr>
<tr>
<td><strong>Rationale</strong></td>
<td>Campbell and colleagues (2014) found that mathematical content and pedagogical knowledge of middle-grades teachers were each directly and positively related to their students’ mathematics achievement.</td>
</tr>
</tbody>
</table>
**Teacher Outcome Construct**

### Fidelity/Integrity

**Definition**
The extent to which a practice (i.e., math curriculum) is delivered as originally developed and specific in plans and protocols.

**Rationale**
Hill and Erickson (2020) found in their review of evidence for fidelity of implementation of STEM curriculums and classroom interventions that better fidelity correlated with better program outcomes. In addition, Hill and Erickson found that classroom observations tended to see more positive fidelity outcomes than teacher self-report. Similarly, Ysseldyke and colleagues (2003) found that math students in classes of high-implementers demonstrated greater academic performance than students in the control group.

### Self-Efficacy: Math Instruction

**Definition**
The degree of confidence in delivering math instruction and meeting student needs.

**Rationale**
Recent research, however, has documented a positive relationship not only between direct measures of teachers' knowledge of mathematics content and pedagogy and student achievement (Baumert et al., 2010; Hill, Rowan, & Ball, 2005) but also between teachers' beliefs about mathematics teaching and learning and student achievement (Love & Kruger, 2005; Clark, 2014).

Teacher self-efficacy has been associated with positive outcomes for both teachers and students. In terms of teacher outcomes, it is a significant predictor of a teacher's understanding of students (Caprara, Barbaranelli, Steca, & Malone, 2006), is associated with a teacher's ability to more effectively respond to students' needs (Hoy & Spero, 2005) and is associated with the development of meaningful curriculum and learning opportunities in the classroom (Reyes, Brackett, Rivers, White, & Salovey, 2012).

### Teacher Self-Efficacy: Cultural Pedagogy

**Definition**
The degree of confidence in using culturally pedagogy strategies.

**Rationale**
Studies indicate that students make academic progress when their teachers are prepared to incorporate culturally relevant teaching strategies into classroom pedagogy (Cammarota & Romero, 2011), and that when teachers recognize and acknowledge students' linguistic and cultural backgrounds and incorporate them into learning, students experience greater academic success (Bui & Fagan, 2013; Dee & Penner, 2017).
THEN Students who are Black, Latino/a, and/or affected by poverty will have improved math student achievement, enjoyment, self-efficacy, and growth mindset.

The EIC Theory of Action acknowledges the complex nature of student-centered learning. Many system, community, school, teacher, family, and student factors influence what students believe about their own ability to do mathematics, how they engage in mathematics coursework, and ultimately, how they perform on local and state benchmark assessments.
Student Outcome Construct
**Beliefs & Self-Efficacy**

**Definition**
For beliefs and self-efficacy, students’ identity, confidence, and understanding that productive struggle and mistakes are opportunities for learning and growing their mind will be examined.

In addition, students’ experience of safety, respect, belonging, inclusion, and joy in the classroom/math learning will be examined.

**Rationale**
Student mathematics self-efficacy and growth mindset has been shown to impact student mathematics performance and enrollment in more rigorous mathematics classes (Evans, 2015; Warwick, 2008; Yeager et al., 2019). In addition, the level of mathematics directly impacts math anxiety and growth mindset (Huang et al., 2019). Long-term, STEM identity mediates the effect of STEM mindset and career interest (Cribbs et al., 2021).

---

**TABLE 4.**
Student Outcome Constructs’ Definitions & Rationales

<table>
<thead>
<tr>
<th>Student Outcome Construct</th>
<th>Definition</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beliefs &amp; Self-Efficacy</strong></td>
<td>For beliefs and self-efficacy, students’ identity, confidence, and understanding that productive struggle and mistakes are opportunities for learning and growing their mind will be examined. In addition, students’ experience of safety, respect, belonging, inclusion, and joy in the classroom/math learning will be examined.</td>
<td>Student mathematics self-efficacy and growth mindset has been shown to impact student mathematics performance and enrollment in more rigorous mathematics classes (Evans, 2015; Warwick, 2008; Yeager et al., 2019). In addition, the level of mathematics directly impacts math anxiety and growth mindset (Huang et al., 2019). Long-term, STEM identity mediates the effect of STEM mindset and career interest (Cribbs et al., 2021).</td>
</tr>
<tr>
<td><strong>Engagement</strong></td>
<td>Engagement is defined as the level of motivation and participation that students display while learning mathematics.</td>
<td>Defined through cognitive, academic, behavioral, and affective engagement, significant evidence exists linking student engagement to academic outcomes (Finn &amp; Zimmer, 2012; Fredricks &amp; McColskey, 2012). Furthermore, existing achievement gaps between children of color and white children are often thought of being a product of difference in engagement - a conclusion not supported by the literature (Bingham &amp; Okagaki, 2012). Skilling and others (2015) highlight the need for attending to student engagement for students in middle school mathematics classes. Less is known about how implementation fidelity/integrity and teacher self-efficacy influence student mathematics engagement.</td>
</tr>
<tr>
<td><strong>Math Achievement</strong></td>
<td>Math achievement is defined as the extent to which students develop a deep knowledge of math and achieve educational goals/standards.</td>
<td>Math achievement can be influenced by a number of different factors. For example, Crawford (2012) found that fidelity of structure (intervention specific fidelity) had a positive effect on student mathematics achievement for middle school students. Additionally, students attending schools with high levels of professional learning activities exhibited more growth than those attending school with lower levels of professional learning activities (Allensworth et al., 2021).</td>
</tr>
</tbody>
</table>
In addition to improving LEA and Provider capacity, the EIC has the goal of producing reliable, practical evidence and measures to inform planning and implementing a district-wide, high-impact math improvement initiative.

The evidence and measures will support:

- **Providers to design more effective implementation services**
- **Districts as they track implementation progress**
- **The field to better understand how high-quality implementation affects enactment, and, in turn, student learning**

A number of public goods (i.e., outputs) will be generated from the EIC for use by the field of education and in other BMGF investments related to the implementation of a high-quality middle years mathematics curriculum within a Coherent Instructional System. See table and figure below for a timeline of the outputs to be generated.
### AT A GLANCE LEARNING QUESTIONS

The At a Glance table for Learning Questions provides a snapshot of what learning questions will be answered and when within the three years of implementation.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Fall Formative</th>
<th>Winter Formative</th>
<th>Spring Formative</th>
<th>Spring Annual Summative</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022-2023</td>
<td>Measurement</td>
<td>Measurement</td>
<td>Measurement</td>
<td>Measurement</td>
</tr>
<tr>
<td>2023-2024</td>
<td>Variability in implementation</td>
<td>Variability in implementation</td>
<td>Application of Established Practice</td>
<td></td>
</tr>
<tr>
<td>2024 December Final Report</td>
<td></td>
<td></td>
<td></td>
<td>Measurement</td>
</tr>
</tbody>
</table>

### AT A GLANCE OUTPUTS

The At a Glance table for Outputs provides a snapshot of what outputs will be produced and when within the three years of implementation.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Fall Formative</th>
<th>Winter Formative</th>
<th>Spring Formative</th>
<th>Spring Annual Summative</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022-2023</td>
<td>Application of Established Practice</td>
<td>Measurement</td>
<td>Measurement</td>
<td>Measurement</td>
</tr>
<tr>
<td>2023-2024</td>
<td>Application of Established Practice</td>
<td>Variability in implementation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024 December Final Report</td>
<td></td>
<td></td>
<td>Variability in implementation</td>
<td>Measurement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Application of Established Practice</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Enabling Context</td>
<td></td>
</tr>
</tbody>
</table>
### 2021-2022 ACADEMIC YEAR

<table>
<thead>
<tr>
<th>Learning questions</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>What conditions most distinguish districts' and schools' readiness for, approach to, and success with implementation? How have districts' priorities for implementation of new curricula or instructional systems shifted in light of COVID-19? How do these shifts influence key enabling conditions?</td>
<td>Definition of readiness conditions at school and district levels that support scaling effective solutions</td>
</tr>
<tr>
<td>Which conditions are go/no-go factors versus important readiness conditions that a provider can help establish?</td>
<td>A curated list of self-assessment tools for district leaders to gauge their system's readiness and identify next steps to prepare for implementation</td>
</tr>
<tr>
<td>What conditions maximize likelihood of successful implementation across contexts relevant to priority students?</td>
<td>Identification of context-specific readiness factors that leaders in priority contexts should consider</td>
</tr>
<tr>
<td>All questions above.</td>
<td>Recommendations for funders, intermediaries, and districts on how to support rebuilding and/or enhancing readiness conditions that have been affected by COVID-19</td>
</tr>
</tbody>
</table>

### 2022-2023 ACADEMIC YEAR

<table>
<thead>
<tr>
<th>Learning questions</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the best-established measures for assessing readiness to take on implementation, self-assessment of what is needed to support implementation, and factors (e.g., system conditions) that should be tracked to support implementation?</td>
<td>Description of the process district leaders use to decide whether and when to implement an evidence-based practice (e.g., based on educator workflows / jobs-to-be-done)</td>
</tr>
<tr>
<td>How do grantees use or eschew established leading and lagging measures related to both CIS and implementation to identify challenges and make improvements to systems of support? Why?</td>
<td>Guidance for how districts can select an implementation support provider based on core and context-specific implementation approaches</td>
</tr>
<tr>
<td>What factors not identified in established measures are critical to identifying challenges and making improvements? How do districts and school systems identify these factors?</td>
<td>Rating of the measurement tools according to educators’ perceptions of their usability and trustworthiness</td>
</tr>
<tr>
<td>All questions above.</td>
<td>Descriptive analysis of why district and school leaders use or eschew data in their decision-making</td>
</tr>
</tbody>
</table>
2023-2024 Academic Year

**Learning questions**

- What is the “menu” of interventions taken by systems to support implementation? Are some seen as more promising and likely to be taken up than others? Which seem to achieve successful implementation most often? How does this vary by context?

- What are the challenges for implementation faced across grantees? What patterns are there to those challenges, especially as related to contextual factors (e.g., student composition, characteristics of previous instructional system, teacher characteristics, broader enabling conditions)? What best practices emerge across the dimensions of implementation support that lead partners are being asked to provide, as well as any others that arise? How are those best practices related to patterns in contextual factors as described above?

- Which lessons learned from implementation research (e.g., curriculum enactment studies, education and health implementation studies) scale across this set of partnerships? Which do not?

- What do educators and system actors identify as the reasons why evidence-based ideas scale or do not scale?

**Outputs**

- An analysis of whether and how implementing a curriculum differs from implementing technical, or smaller scale solutions (e.g., a new medicine or a direct-to-student EdTech solution)

- Identification of core versus context-specific approaches to curriculum implementation

- A suite of measurement tools and new evidence to inform design and progress monitoring of solution implementation at scale
<table>
<thead>
<tr>
<th>VARIABILITY IN IMPLEMENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the “menu” of interventions taken by systems to support implementation? Are some seen as more promising and likely to be taken up than others? Which seem to achieve successful implementation most often? How does this vary by context?</td>
</tr>
<tr>
<td>An analysis of whether and how implementing a curriculum differs from implementing technical, or smaller scale solutions (e.g., a new medicine or a direct-to-student EdTech solution)</td>
</tr>
<tr>
<td>What are the challenges for implementation faced across grantees? What patterns are there to those challenges, especially as related to contextual factors (e.g., student composition, characteristics of previous instructional system, teacher characteristics, broader enabling conditions)?</td>
</tr>
<tr>
<td>What best practices emerge across the dimensions of implementation support that lead partners are being asked to provide, as well as any others that arise? How are those best practices related to patterns in contextual factors as described above?</td>
</tr>
<tr>
<td>Identification of core versus context-specific approaches to curriculum implementation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPLICATION OF ESTABLISHED PRACTICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which lessons learned from implementation research (e.g., curriculum enactment studies, education and health implementation studies) scale across this set of partnerships? Which do not?</td>
</tr>
<tr>
<td>What do educators and system actors identify as the reasons why evidence-based ideas scale or do not scale?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MEASUREMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>What defines effective implementation? What are evidence-based leading indicators, lagging indicators, and data sources to assess progress and impact?</td>
</tr>
<tr>
<td>How, and how much, did implementation affect student math learning and/or leading indicators of math learning? In what contexts and for whom?</td>
</tr>
<tr>
<td>How do the findings from this cohort reinforce, complement, or differ from the findings about classroom-level factors that drive curricula’s effectiveness in the separate enactment studies that the Foundation is funding?</td>
</tr>
<tr>
<td>How does teacher perception (e.g., feasibility, acceptability, and appropriateness) of the math curriculum within CIS influence implementation?</td>
</tr>
</tbody>
</table>

| A definition of effective implementation that districts and researchers can use to describe the strength of an implementation process |

| A curriculum-specific model of active implementation that highlights curriculum-specific emphases such as an implementation stage based approach for curriculum |

<table>
<thead>
<tr>
<th>DETAILED TIMELINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2024 December Final Report</td>
</tr>
<tr>
<td>What defines effective implementation? What are evidence-based leading indicators, lagging indicators, and data sources to assess progress and impact?</td>
</tr>
<tr>
<td>How, and how much, did implementation affect student math learning and/or leading indicators of math learning? In what contexts and for whom?</td>
</tr>
<tr>
<td>How do the findings from this cohort reinforce, complement, or differ from the findings about classroom-level factors that drive curricula’s effectiveness in the separate enactment studies that the Foundation is funding?</td>
</tr>
<tr>
<td>How does teacher perception (e.g., feasibility, acceptability, and appropriateness) of the math curriculum within CIS influence implementation?</td>
</tr>
<tr>
<td>A definition of effective implementation that districts and researchers can use to describe the strength of an implementation process</td>
</tr>
</tbody>
</table>

| A curriculum-specific model of active implementation that highlights curriculum-specific emphases such as an implementation stage based approach for curriculum |
SAMPLE

Participants for this investment include 10 providers and 19 Local Education Agencies (LEAs) across the United States.

Each provider has partnered with 2 to 3 LEAs for a total of 19 Provider-LEA dyads.

Criteria for inclusion in this investment included selection of a high-quality middle years mathematics curriculum, established success in providing effective professional learning services for successful implementation of the curriculum, representation of the priority student population, and ability to develop and execute data sharing agreement for each dyad.

Across LEAs, six different high-quality math curricula have been identified, including: Go Math (Texas) (n =3), Eureka Math (n =1), Ready Math (n=2), Illustrative Math (n=8), Engage NY (n=1), Open up Resource (n=1), Carnegie Learning (n=1), and Agile Mind (n=2).
Multiple data collection methods were used in Phase I to examine pre-implementation conditions and to inform progress monitoring. Readiness is a condition that needs to be developed and maintained throughout implementation. Given this, NIRN will continue to monitor readiness indicators and will also collect data and information from Providers and LEA partnerships in Phase II to answer the Cohort-wide learning questions.

Data will be collected through observations, electronic surveys, semi-structured interviews, and collection of partnership-specific learning agenda data and implementation plans.

*Table 5* (p.35) depicts the data collection timeline for Phase I.
OBSERVATION MEASURES

Two observation measures will be used to collect data focused on the readiness constructs of District Teaming, LEA Executive Sponsor Engagement, Communication, Assessing Fit & Feasibility, and Implementation Planning. Implementation Team LookFors (Cusumano, Preston, & Ward, 2017) will be used to collect data on the quality of tools and resources developed by the Provider-LEA partnerships. A Support Log (NIRN, 2021) will be used to collect data on session dosage (with NIRN), topics covered, identify barriers, facilitators and successes. Data using these two observation measures will be collected at every Provider-Dyad session but only 3 data points was used in the analysis (i.e., March, April, and May) in Phase I.

INTERVIEWS

Semi-structured interviews will be conducted with providers and participating LEA executive sponsors to collect data on the readiness constructs of District Teaming, LEA Executive Sponsor Engagement, Communication, Assessing Fit & Feasibility, and Implementation Planning. In Phase I, the semi-structured interviews were conducted by the end of May 2021.

IMPLEMENTATION TEAM (IT) SURVEY

The IT survey is a self-report survey which consists of 18 items. This survey will be used to collect data on the readiness constructs of District Teaming, LEA Executive Sponsor Engagement, Communication, Assessing Fit & Feasibility, and Implementation Planning. The survey was administered electronically through Qualtrics survey software once during Phase I to those with a team formed.

PERMANENT PRODUCTS REVIEW

A product review rubric was used to review the implementation plan. These data will be centered on the readiness constructs of Communication, Assessing Fit & Feasibility, Implementation Planning, and Measurement Planning. Implementation specialists assigned to each Provider-LEA partnership reviewed the implementation plan at least twice in Phase I.
### TABLE 5.
Phase I Data Collection Measures and Timeline

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>DATA COLLECTION METHOD</th>
<th>FREQUENCY &amp; SCHEDULE</th>
<th>WHO WILL COLLECT?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation Team Survey</td>
<td>Online Survey</td>
<td>1x (March-June)</td>
<td>EIC Team</td>
</tr>
<tr>
<td>Team LookFors checklist</td>
<td>Observation</td>
<td>2x (March-May)</td>
<td>EIC Team</td>
</tr>
<tr>
<td>Support Log</td>
<td>Observation</td>
<td>3x (May/June)</td>
<td>EIC Team</td>
</tr>
<tr>
<td>Provider Interviews</td>
<td>Semi-structured Interview</td>
<td>1x April - May</td>
<td>EIC Team</td>
</tr>
<tr>
<td>Exec. Sponsor Interviews</td>
<td>Semi-structured interview</td>
<td>1x April - May</td>
<td>EIC Team</td>
</tr>
<tr>
<td>Permanent Products Rubric</td>
<td>Product Review (Communication plan,</td>
<td>1x May</td>
<td>EIC Team</td>
</tr>
<tr>
<td></td>
<td>implementation plan, and fit and feasibility)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### MEASURES
Implementation, Teacher, & Student Outcome Measures

During Phase II, a number of implementation, teacher, and student outcome constructs will be examined for delivery of professional learning services and the implementation of high-quality middle years mathematics curriculum. These are organized into short-term and intermediate implementation outcomes for school and district leadership and teams, and long-term outcomes for teachers and students. The readiness indicator measures in Phase I will continue to be collected in addition to these implementation, teacher, and student outcome measures (See Figure 3 p.36).
IMPLEMENTATION OUTCOMES (SHORT TERM & INTERMEDIATE)

Within the EIC, high-quality professional learning services and a high-quality middle years mathematics curricula are being implemented. Distinguishing implementation effectiveness from “treatment” effectiveness is necessary for translating evidence into use within classrooms. It is important to be able to determine whether lack of progress in improving outcomes for students and teachers occurred because the “treatment” was ineffective in the setting or whether the “treatment” was deployed incorrectly (i.e., implementation failure). Proctor et al. (2011) defined implementation outcomes as “the effects of deliberate and purposive actions to implement new treatments, practices, or services.” Implementation outcomes serve as indicators of implementation progress and serve as key intermediate outcomes in relation to student/teacher outcomes. Specifically, implementation outcomes are the necessary pre-conditions for obtaining desired changes in teacher and student outcomes. Conceptualizing and measuring these implementation outcomes (see Figure 3) in Phase II of the EIC will help advance understanding of implementation processes, enable comparative effectiveness of implementation strategies, and enhance efficiency in implementation research (Proctor, 2020).

To measure these implementation outcomes, a variety of methods are proposed including collecting administrative data, observations, interviews, and surveys.
A teacher survey will be used to measure the outcomes of Acceptability, Appropriateness, and Feasibility. The survey will be administered by the Provider and/or LEA once per year in Phase II and shared annually with the Learning Partner (NIRN). The following measures have been shared with Providers and LEAs to inform their survey development.

Acceptability, or the perception of satisfaction with the professional learning services and math curriculum, will be measured using the *Acceptability of Intervention Measure* (Weiner et al, 2017). The Acceptability of Intervention Measure consists of five items on a 5-point Likert scale from “completely disagree” to “completely agree.”

Appropriateness will be measured using learnings from the *BMGF’s PLP investment*. Specifically, for appropriateness of professional learning services, survey items from the following scales were shared: the Professional Learning Authority Scale, Professional Learning Consistency Scale, and Professional Learning Specificity Scale. Appropriateness of the high-quality math curriculum items were shared from the Curriculum Consistency Scale and the Appropriateness of Intervention Measure (Weiner et al., 2017).

**PROFESSIONAL LEARNING SPECIFICITY SCALE**
measures degree of detail and clarity regarding areas such as objectives and learning goals, and connection to curriculum and state standards. It consists of six items on a 6-point Likert scale from “completely disagree” to “completely agree.”

**PROFESSIONAL LEARNING CONSISTENCY SCALE**
measures the degree of alignment between professional learning activities, the school’s mission and goals, and the district’s policies. It consists of five items on a 6-point Likert scale from “completely disagree” to “completely agree.”

**PROFESSIONAL LEARNING AUTHORITY SCALE**
measures the extent to which teachers buy into and feel supported in implementing the strategies taught in professional learning. It consists of four items on a 6-point Likert scale from “completely disagree” to “completely agree.”
CURRICULUM CONSISTENCY SCALE
measures the degree of alignment between the curriculum and areas such as content standards, assessments, and school/district policies. It consists of seven items on a 6-point Likert scale from “completely disagree” to “completely agree.”

APPROPRIATENESS OF INTERVENTION MEASURE
(Weiner et al., 2017) measures the perceived fit, relevance, or compatibility of the curriculum and consists of five items on a 5-point Likert scale from “completely disagree” to “completely agree.”

Feasibility of the professional learning services and high-quality middle years math curriculum will be measured using the Feasibility of Intervention Measure (Weiner et al., 2017), which consists of five items on a 5-point Likert scale from “completely disagree” to “completely agree.”

Quality of Professional Learning Services: Coaching will be measured through surveys. Items from the Relationship with Coach scale and the Coaching Satisfaction Survey (Ihlo et al., 2017) were shared to inform survey development.

RELATIONSHIP WITH COACH SCALE
measures the degree of trust and respect established between the teacher and coach and consists of six items on a 4-point Likert scale of “not at all” to “a great extent.”

COACHING SATISFACTION SURVEY
(Ihlo et al., 2017) which measures the perception of the coaching services quality and satisfaction on a 4-point Likert scale of “completely disagree” to “completely agree.”
PRINCIPAL SURVEY

Principals will be surveyed by the Learning Partner (NIRN) to measure their leadership for implementation and implementation outcomes of Acceptability, Appropriateness, and Feasibility. The principal survey will be administered annually in Phase II. The survey will consist of items from the following measures: Implementation Leadership Scale (Aarons et al., 2014), Acceptability of Intervention Measure, Feasibility of Intervention Measure, and Appropriateness of Intervention Measure (Weiner et al., 2017). The survey will be administered by the Learning Partner (NIRN) once per year and will be shared with Providers and LEAs.

IMPLEMENTATION TEAM SURVEY

Members of the District Implementation Team will be surveyed by the Learner Partner (NIRN) to measure the implementation outcome of Organizational Capacity. Survey items (n = 18) will measure the functioning of the team, available leadership supports, and use of data. The survey will be administered annually by the Learning Partner (NIRN) and shared with Providers and LEAs. Note: This is the same survey that was used in Phase I.
DISTRICT CAPACITY ASSESSMENT

The District Capacity Assessment (DCA; Ward et al., 2015) will collect data about roles, structures, functions, and systemic infrastructure to assess the district capacity to implement a high-quality middle grade math curriculum. Specifically, the DCA will collect data focused on the readiness constructs of District Teaming, Implementation Planning, and Measurement Planning. The DCA consists of 26-items scored on a three-point scale (i.e., 2 – Fully in Place, 1 – Partially in Place, 0 – Not in Place). The assessment is completed by a trained administrator with a facilitator and team of LEA/provider participants. A total score and three subscale scores (i.e., Competency, Leadership, and Decision Support Data System) are generated. The DCA is designed to guide LEA teams to align implementation efforts and resources for the selected high-quality middle years mathematics curriculum. Psychometrically, the DCA’s content validity has been established. It has an adequate internal structure (RMSEA = .071, CFI = .93, TLI = .92), internal consistency (Cronbach alphas of 0.91 for the total score and 0.79 to 0.81 for the subscale scores), and test-retest reliability (r = .98 for Leadership, .78 for Decision Support Data System and Competency Scales; Ward et al., 2020). The DCA will be administered annually by the Learning Partner (NIRN) with the district implementation team in Phase II and shared with the Provider and LEA.

INTERVIEWS

Semi-structured interviews will be conducted by the Learning Partner (NIRN) with providers and participating LEA executive sponsors, and coaches to collect data on the implementation outcomes of organizational capacity, quality of professional learning services and implementation, and sustainability. Sustainability questions will be drawn from the Program Sustainability Tool (Washington University, 2018). Interviews will be conducted annually in Phase II by the Learning Partner (NIRN).

ADMINISTRATIVE DATA

Administrative data will be submitted annually by the LEA to measure the outcomes of Adoption, Cost, Penetration/Reach and Sustainability. Administrative data will be shared with the Learning Partner (NIRN) annually.

*Adoption and Penetration/Reach* will be measured by the number of schools and teachers using High-Quality Mathematics Curriculum.

*Cost* will be measured by review of resources allocated on the implementation plans.
The intermediate teacher outcomes of self-efficacy, or degree of confidence in delivering math instruction, and using cultural pedagogy strategies, as well as fidelity/integrity of the math curriculum will be measured in Phase II (see Figure 3 p.36) by Providers and LEAs. The data collected will then be shared with the Learner Partner (NIRN). Table 6 (p.42) depicts the data collection timeline for all short-term and intermediate outcomes for Phase II.

**TEACHER SURVEY**

A teacher survey will be used to measure the degree of confidence in delivering math instruction, meeting student needs, and using culturally pedagogy strategies. Using learnings from the BMGF’s PLP investment, survey items will be shared with Providers-LEAs to use from Confidence in Meeting Students Needs Scale, Culturally Responsive Teaching Self-efficacy Scale, and Confidence Using Ambitious Math Instruction Scale. In addition, Providers will identify instruments to assess teacher growth in knowledge and skill in delivery of mathematics instruction aligned with the professional learning being provided. The teacher survey will be administered once a year in Phase II by the Provider or LEA and shared with the Learning Partner (NIRN) annually.

**CONFIDENCE IN MEETING STUDENTS NEEDS SCALE**

consists of six items on a 10-point Likert scale from 0 (not confident at all) to 10 (extremely confident).

**CULTURALLY RESPONSIVE TEACHING SELF-EFFICACY SCALE**

consists of eight items on a 10-point Likert scale from 0 (not confident at all) to 10 (extremely confident).

**CONFIDENCE USING AMBITIOUS MATH INSTRUCTION SCALE**

consists of eight items on a 10-point Likert scale from 0 (not confident at all) to 10 (extremely confident).

**CLASSROOM OBSERVATIONS**

Classroom observations of use of the curriculum as intended (integrity/fidelity) will be conducted by Providers and LEA staff using observation tools identified by the Provider and the LEA. Examples of observation tools proposed by Providers include those such as MQI Observation Tool, EQUIP tool (Reinholz and Shah, 2018), Illustrative Mathematics’ IM Teams, and Classroom Observation Tool. Observation data collected by the Provider and LEA will be shared annually with the Learning Partner (NIRN).
### TABLE 6.
Phase II Implementation Outcomes Data Collection, Measures and Timeline

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>DATA COLLECTION METHOD</th>
<th>WHO WILL COLLECT?</th>
<th>FREQUENCY &amp; SCHEDULE</th>
<th>OUTCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative Data Form for Reporting</td>
<td>Administrative Data</td>
<td>Provider/LEA</td>
<td>Annually Spring</td>
<td>AD, P/R, SU</td>
</tr>
<tr>
<td>District Capacity Assessment (DCA)</td>
<td>District Team Assessment</td>
<td>EIC Team</td>
<td>Annually Winter</td>
<td>OC</td>
</tr>
<tr>
<td>Implementation Team Survey</td>
<td>District Team Survey</td>
<td>Provider/LEA</td>
<td>Year 1: Bi-annually Summer-Winter Year 2-3: Annually Winter</td>
<td>QU, OC, CO</td>
</tr>
<tr>
<td>Implementation Leadership Scale</td>
<td>Principal Survey</td>
<td>Provider/LEA</td>
<td>Annually - Spring</td>
<td>QU, OC</td>
</tr>
<tr>
<td>Measure to be selected by Providers &amp; LEAs using surveys above (see example teacher survey)</td>
<td>Teacher Survey</td>
<td>Provider/LEA</td>
<td>Annually - Winter</td>
<td>AC, AD, AP, FE, QU, OC, TE</td>
</tr>
<tr>
<td>EIC DT and Leadership Interview Protocol</td>
<td>District team (DT) and Leadership Interviews</td>
<td>Learning Partner</td>
<td>Annually - Winter</td>
<td>OC, SU</td>
</tr>
<tr>
<td>EIC Coach Interview Protocol</td>
<td>Coach Interviews</td>
<td>Learning Partner</td>
<td>Annually - Winter</td>
<td>QU, OC, SU</td>
</tr>
<tr>
<td>Implementation Team LookFors checklist</td>
<td>Observation</td>
<td>Learning Partner</td>
<td>Annually - Winter</td>
<td>QU, OC</td>
</tr>
<tr>
<td>Observation of Math Instruction (Classroom)</td>
<td>Observation</td>
<td>Provider/LEA</td>
<td>Annually - Winter</td>
<td>QU</td>
</tr>
<tr>
<td>Implementation Plan Review Rubric</td>
<td>Product Review</td>
<td>Learning Partner</td>
<td>Annually - Spring</td>
<td>CO</td>
</tr>
</tbody>
</table>

Note. AC = Acceptability; AD = Adoption; AP = Appropriateness; FE = Feasibility; QU = Quality of PL Services; OC = Organizational Capacity; CO = Cost; I/FI = Integrity/Fidelity; P/R = Penetration/Reach; SU = Sustainability
DEMOGRAPHIC DATA

District and School-level variables will be compiled to provide context for evaluation data including number of students enrolled, student-teacher ratio, percentage of students receiving free or reduced lunch, percentage of students in each racial/ethnic subgroup, and home language. (See Appendix A for a complete list of school, student, and teacher level demographic data). These data will be provided annually by the Provider/LEA. Socio-demographic data will be described in tables and depicted in charts.

STUDENT OUTCOMES (LONG-TERM)

Student outcome measures will be examined across the domains of student math achievement, experience in the classroom, and engagement. Table 7 (p.44) depicts the data collection timeline for student outcomes in Phase II.

STUDENT MATH ACHIEVEMENT MEASURES

To measure student math achievement, state and local assessments (e.g., benchmark assessments) will be used. This data will be collected by all LEAs and shared with the Learning Partner annually. For the group of LEAs that are also participating in the AMS investment funded by BMGF (n = 4 LEAs potentially), student work will also be shared with the Learning Partner in Year 3 only to be analyzed using methods such as Stein & Smith’s Math Task Analysis Guide or the Math Assignment Analysis Guide. The analysis method is being determined with Mathematica, AMS Primary Investigator.

STUDENT EXPERIENCE, BELIEFS, AND ENGAGEMENT RELATED TO MATHEMATICS

To measure students’ self-efficacy and confidence, value of math, enjoyment of math, and engagement with math, a student survey will be used by all LEAs. Using recommendations from BMGF Middle Years Math Cohort, a short student survey (n = 28 items) was developed using scales and items from the Expectancy-Cost-Value Scale (Kosovich et al., 2014; Lauermann et al., 2017; Simpkins et al., 2006; Wang, 2012), Math and Me Survey (Adelson & McCoach, 2011), Math and Science Engagement Scales (Wang et al., 2016; Fredricks et al., 2016), National Study of Learning Mindsets (Yeager, et al., 2019), and Trends in International Mathematics and Science Survey (TIMSS 2011). The student survey will be administered by all LEAs annually and shared with the Learning Partner (NIRN) annually.
### TABLE 7.
Student Data Collection, Measures and Timeline

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>DATA COLLECTION METHOD?</th>
<th>WHO WILL COLLECT?</th>
<th>FREQUENCY &amp; SCHEDULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adapted AMS Survey</td>
<td>Student Survey</td>
<td>Provider/LEA</td>
<td>Annually - Spring</td>
</tr>
<tr>
<td>Student administrative records of state and local benchmark assessments</td>
<td>Student Administrative Records (e.g., student math achievement data,) See Appendix A for list of variables</td>
<td>Provider/LEA</td>
<td>Annually - Fall</td>
</tr>
<tr>
<td><strong>Stein &amp; Smith’s Math Task Analysis Guide</strong> or the Math Assignment Analysis Guide</td>
<td>Student Work (subsample)</td>
<td>Provider/LEA</td>
<td>Spring (Year 3 only)</td>
</tr>
</tbody>
</table>

**MEASURES**  
**EIC Learning Partner Effectiveness Measures**

To evaluate the effectiveness of the Learning Partner’s (NIRN) ability to provide high-quality services and tools, a number of formative and summative measures will be used, including: (1) professional learning evaluations to capture the quality and use of adult learning practices, knowledge gained, and follow up learning needs at every learning event; (2) consultation/coaching support effectiveness surveys and documentation logs will be monitored every three months; and (3) tracking of iterative development process from draft to finalization as compared to identified milestones and benchmarks every three months.

To evaluate the effectiveness of NIRN’s ability to co-design and execute a cohort-wide learning plan inclusive of measures and baseline data, formative and summative measures will be used, including: (1) evaluation surveys to assess quality and effectiveness of facilitation and support in the co-design process every three months; (2) professional learning evaluations to capture the quality and use of adult learning practices, knowledge gained, and follow up learning needs at every learning event; and (3) formative feedback gathered throughout the process through observations.
Effective Implementation Cohort (EIC) Cohort Wide Learning Plan

Providers self-selected their partnerships with the LEAs to participate in this investment. LEAs are selecting the participating schools within their district during Phase I. The EIC team will hold an overview of the investment at Provider-LEA partnership sessions. The overview will include a review of the purpose, procedures, what and how data will be collected, potential risks, and their right to withdraw from the data collection activities at any time. In addition, as part of our role as a Learning Partner, cohort wide support sessions will be held quarterly to support providers and LEA teams in their use of implementation best practices and use of data for continuous improvement purposes.

This evaluation entails data collection from multiple sources, including online systems, teacher surveys, school records, and observational data collection. When possible, we will use web-based data collection to improve efficiency. Tables 5 (p.35) and Table 7 (p.44) above show the data collection timeline.

## STUDENTS

### MATH ACHIEVEMENT ASSESSMENTS

LEAs will share data from math student achievement assessments that students participate in as part of their typical educational experience. Specifically, math scale score data from state summative assessments (completed annually) and their local selected benchmark assessments (e.g., NWEA Measures of Academic Progress, i-Ready, STAR; completed three times a year) will be shared with the research team annually (Fall: October-November) for previous year assessment results.

### SURVEY

LEAs/providers will share data from a web-based student survey regarding student’s beliefs, engagement, and experience with math instruction. The student survey will be administered by the LEA or provider annually in the Spring (April - June) and shared annually with the Learning Partner (NIRN).
**MATH TEACHERS**

**SURVEY**

Math teachers within participating schools will complete a web-based survey administered by their LEA or Provider partner annually (Winter: December-February) to assess their perception of the implementation of the math curriculum, professional learning services (e.g., feasibility, appropriateness, acceptability, quality), knowledge and skill in delivery of instruction, and their confidence in teaching mathematics and using culturally responsive teaching practices. The LEA or provider will share survey results annually with the Learning Partner (NIRN).

**OBSERVATION**

Math teachers will experience observations of their delivery of instruction by a trained observer who are either LEA (e.g., instructional coach) or Provider staff (e.g., trainer, instructional coach). Providers and LEAs will conduct observations for 30 minutes using their selected observation protocol annually in the Winter (January-February) and share the data annually with the Learning Partner (NIRN). Observers are trained by the Provider (i.e., external partner providing support to the district) and will be required to meet and maintain 80% or higher rate of agreement with the trainer in project classrooms. Reliability will be periodically tested and reported by the provider to the research team to detect observer drift and maintain consistency.

---

**SCHOOL PRINCIPALS**

**SURVEY**

Principals within participating schools will complete a web-based survey (*Implementation Leadership Scale*, 12 items on a 5-point Likert Scale), administered by the Learning Partner (NIRN) annually (April-June) to assess their leadership practices within implementation.

**PROVIDER**

**INTERVIEWS**

Semi-structured interviews will be conducted with the lead Provider for the LEA (e.g., external partner to the LEA providing professional learning supports). All interviews will be conducted annually (Winter: December-February) for 30-45 minutes by the research team. Interviews will be conducted virtually and recorded with permission from the participant for transcription and analyses.
DISTRICT STAFF

INTERVIEWS
Semi-structured interviews will be conducted with a district executive leader (e.g., Chief Academic Officer, Assistant Superintendent). Interviews will also be conducted with instructional coaches providing instructional coaching to teachers, if available. All interviews will be conducted annually (Winter: December-February) for 30-45 minutes by the research team. Interviews will be conducted virtually and recorded with permission for transcription and analyses.

SURVEY
District staff (1) participating on the district implementation team accountable for math implementation and (2) district staff serving in the role of instructional coaches will complete a self-report web-based survey administered by the research team annually (Winter: December - February) to assess their leadership practices within implementation. The survey will be 18 questions on a 4-point Likert scale.

CAPACITY ASSESSMENT
District staff participating on the district implementation team accountable for math implementation will complete a team-based district capacity assessment. The team assessment will be administered by a trained facilitator on the research team using a consensus-based scoring procedure. The assessment takes 90-120 minutes to complete. The research team will administer the team capacity assessment annually (Winter: December-February).

OBSERVATIONS
The district implementation team will be observed in their implementation team meetings by a trained observer on the research team. The trained observer from the research team will observe team meetings annually for 30-60 minutes in length (Winter: December-February).
Data analysis will be conducted to answer the learning questions. We will use a mixed-method approach to address each category of questions through an analysis of qualitative and quantitative data sources. The analysis will be largely descriptive using cohort-wide trends to achieve goals including:

Creating evidenced-based models for the implementation of high-quality middle years math curricula

Adding to the knowledge base available to education professions on how to implement high-quality middle years math curricula, particularly for students of color and those experiencing poverty

Providing participating Providers and LEAs with meaningful, formative feedback to inform their practice.

The design will use the school-level as the primary unit of analysis to examine all the learning questions. Analysis of the qualitative and quantitative data will be through descriptive, visual trends, and thematic analyses. Analytically, the challenge of this investment is that there are few common measures across all schools. Within the sample, not all schools use the same high-quality math curriculum. Within a given curriculum, the sample might be smaller. Small numbers raise the risk that observed changes or trends could be due to random variation in these small groups. To the extent possible we will aggregate data across schools and LEAs. When examining math academic outcomes, we will seek to disaggregate data by race, ethnicity, home language, and income level. Finally, a control/comparison group is not being used given that the purpose of this investment is to explore implementation conditions and not establish causal relationships.

To investigate more intensively the learning question of how implementation affects student math learning, we will conduct analyses at the student- and teacher-level as the primary unit of analysis with a select number of districts. We will strive to have a subsample that is representative of key factors/variables including priority student representation and various district context factors such as locale and size.
QUALITATIVE ANALYSES

The data from all interviews will be analyzed qualitatively and summarized in stages. First, all interviews will be recorded and transcribed verbatim. After each interview, NIRN will use the implementation constructs as themes, in addition to developing themes organically as the coding occurs. We will use Thomas’ (2006) general inductive approach to qualitative analysis. This approach provides researchers with a straightforward method to 1) condense raw text into a brief, summary format; 2) establish clear links between the learning questions and the summary findings derived from the raw data; and 3) develop a framework of the underlying structure of experiences or processes that emerge from the raw data. NVivo software will be used for coding and to generate themes based on the key learning questions being asked. Two members of the NIRN team will code the data to ensure inter-rater reliability. After coding separately, the coders will come together to talk about similarities and differences in codes. The themes will be synthesized into the interim and final report.

QUANTITATIVE ANALYSES

MISSING DATA

To account for missing data, we will use multiple imputations in SAS 9.4. All predictor and covariate variables will be included in the imputation procedure. Creating multiple datasets with reasonable missing data values and aggregating results from analyses using multiple datasets provide the best approximation of relationships among variables.

This approach is preferred over listwise deletion or single imputation. We expect limited missing data (<1% student attrition over the year; no classroom attrition expected) but will use multiple imputation using MCMC algorithms to impute missing data (Schafer & Graham, 2002) into 40 imputation datasets. These approaches for addressing missing data are appropriate when data are missing at random (Schafer & Graham, 2002).

STATISTICAL POWER

Power analysis will be conducted to determine the minimal detectable sized effect using the dataset and ensure that the study is well powered to conduct the proposed analyses and to detect the smallest-sized effects.

PRELIMINARY ANALYSES

Preliminary factor analyses will test whether composite variables underlie our direct assessments of student's outcome measures. This will provide more precise measurement and potentially fewer tests of our constructs. For quantitative analyses using multiple measures for the same construct (e.g., teacher surveys of implementation), initial descriptives for items under each construct will be examined in order to conduct initial assessment of variability in item responses prior to psychometric modeling. For continuous items, this will include means, Pearson correlations, and standard deviations; for binary items (e.g., correct/incorrect), this will include item frequencies and tetrachoric correlations. For each construct, formalized psychometric analysis will proceed with the following steps.
First, confirmation of the factor structure (unidimensional or multidimensional, contingent on the structure of the established measure) as it has been used in past studies using a flexible form of confirmatory factor analysis (CFA) that allows for a mix of item distributions (continuous or ordered categorical).

Second, contingent on model fit, a more restrictive form of CFA model will be fit where factor loadings will be constrained to equality within each factor; this has the effect of fitting a model to test whether total scores are appropriate to use (i.e., equal weighting of items; McNeish & Wolf, 2020). The degree of misfit of this model captures the extent to which using total scores would create bias in scale score estimates and give an inaccurate picture of participant progress at the individual- and aggregate-levels. Third, once a finalized base model is established, differential item functioning (DIF) will be tested to assess the extent to which different item parameters (i.e., item intercepts/thresholds, factor loadings) are required across different predictors of DIF such as time or demographic variables (e.g., race/ethnicity) under the moderated non-linear factor analysis (MNLFA; Bauer, 2017) framework; see also Morgan-López et al., 2020a, 2020b; Ruglass et al., 2020; Saavedra et al., 2021.

Scale scores and “personalized” standard errors of measurement that take into account a) differences in the relative weight of each item in relation to the construct and b) whether the items parameters differ across time and/or populations (i.e., free of measurement bias) would then be output from MNLFA analyses. The outputs will then be used to conduct school and student level analyses using Mplus v8.
Standardized statistical approaches will be followed to examine student math achievement data. When standardization is not feasible, harmonization will be explored as an alternative methodology. Logical harmonization involves integrating similar data elements without any data transformation. However, harmonization should involve statistical transformation, usually using Item Response Theory to bring the different data elements to the same metrics. If math assessment outcomes are statistically harmonizable, they will be integrated, and appropriate analyses will be used to create equitable scales. If instruments are not harmonizable, data-based outcomes (e.g., dichotomous variable, improvement/no improvement on study/site specific scale) will be integrated.

SCHOOL-LEVEL UNIT OF ANALYSIS

The school-level analysis will be descriptive in nature and will explore the bi-directional relationship between level of implementation support, aggregate school implementation conditions (average rates of fidelity, teacher self-efficacy, etc.), and aggregate student mathematics beliefs, experience, engagement, and achievement. While teacher and student surveys and administrative data will be collected at the individual-level, it will be aggregated to the school-level because we are not requiring a cohort-wide process of linking teacher and student data.
Through these analyses, we will be able to describe generally the extent to which provider, LEA, and school implementation conditions influence teacher self-efficacy, implementation fidelity, and aggregate student outcomes (especially for groups of priority students); and whether school-level teacher self-efficacy and implementation fidelity is related to positive student outcomes aggregated to the school-level.

STUDENT-LEVEL UNIT OF ANALYSIS

A multitude of district, school, and teacher factors influence not only how well a new curriculum is implemented, but also how students perceive and engage with mathematics and achieve on standardized state and local assessments. While more rigorous and demanding than a school-level analysis, a deeper dive with a selected group of LEAs will allow us to gain a better understanding of how implementation affects student learning, for whom, and in what contexts. In these analyses, student administrative data would be matched to teacher administrative data to control for the potential influence of teacher-level factors on student learning and achievement. We will use propensity score analysis to ensure the group of students in the schools being included in the analysis are similar on key variables (race/ethnicity, gender, economically disadvantaged, academic achievement) as students in schools not included in the analysis. Propensity score analysis (with weights) is used to balance measured confounders or covariates that influence both the probability of selection into two or more non-experimental groups and intervention outcomes. Propensity score analysis is conducted in three steps:

1. each student’s probability of receiving the treatment based on a set of covariates,
2. ensure balance between the two groups on key variables,
3. conduct multivariate analyses using the outcome(s) of interest.

We will use multilevel linear modelling to account for the nested nature of the educational data (students in a classroom within a school, within a LEA). Although there are limitations with regards to generalizability, controlling for school and teacher factors that might influence implementation fidelity and student outcomes will allow to isolate the effect of implementation on student learning, for whom, and in what contexts.

DATA MANAGEMENT

Data management is maintained in accordance with protocols and policies of the LEA and their Institutional Review Board (IRB) or Research Review Board (RRB) and other parties’ IRBs including UNC-Chapel Hill’s for NIRN. In addition, procedures and data sharing will be reviewed for compliance with the Family Educational Rights and Privacy Act (FERPA) and, if required, Health Insurance Portability and Accountability Act (HIPAA).
DATA ACCESS

Research datasets answering the Cohort’s Learning Questions will be made available in common, accessible data formats, such as those for SPSS or SAS software, Excel files, or comma-separated text files. Supporting documentation will be made available in common file formats such as Microsoft Word or Excel, Portable Document Format (PDF), or other commonly available software (e.g., Google Docs).

All data sets will be publicly accessible within 12 months of the completion of the investment and will remain accessible for a minimum of ten (10) years after the end of the project. We will pursue archiving the combined datasets with the UNC Odum Institute’s Dataverse Network (DVN). The final dataset to be shared will be completely de-identified with removal of student, teacher, school, and district names and any other personally identifiable information (PII).

Confidentiality of data will be maintained according to the policies and procedures identified in the human subjects’ protocol that will be approved by UNC-CH’s IRB before beginning research activities for this project. For data use by investment staff, all electronic data will be securely stored in either the investment database or file server, maintained by UNC-Chapel Hill, and accessible only by authorized investment personnel and organization staff members responsible for managing computing equipment. Data exchanged with investment staff outside UNC-Chapel Hill will be encrypted prior to sharing.

Given the variety of data to be collected, the Provider/LEAs will submit the data via a secure file transfer protocol (SFTP) for which each Provider and LEA will be granted unique access credentials to create a separate submission portal. Data storage on the system will adhere to all laws protecting confidential information. The data team will monitor and review the data submitted to ensure high quality and completeness throughout the investment.

The first data upload will include data from the 2020-2021 school year. These data should be shared by October 30, 2021. If the data uploaded is not de-identified, the EIC team at UNC will de-identify the data and summarize before sharing. The data will be reviewed to determine if subjects could be identified deductively through small cells, and any offending demographic variables will be removed to minimize such deductive disclosure for small subsamples.

Semi-structured interview data will be recorded for accuracy and quality of data collection with the permission of participants. All data will be securely stored, and no identifiable information will be stored.

Semi-structured interview data will be recorded for accuracy and quality of data collection with the permission of participants. All data will be securely stored, and no identifiable information will be stored.
REFERENCES


**APPENDIX A**

The data being requested will include records for all middle school students enrolled at any time in the given year requested and attending schools participating in the EIC project. While the EIC is focused on students who are experiencing poverty, Black, Latino/a, and/or English Learner (EL)-Designated (“priority students”), it is critical to have data on all students from all participating schools to address the learning questions of the investment. **All measures should be reported by student once a year.** A basic description of the variables to be requested appears in the table below.

All student data should be shared for the following academic years: 2020-2021, 2021-2022, 2022-2023, 2023-2024.

Teacher data should be shared for the following academic years: 2021-2022, 2022-2023, 2023-2024.

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>VARIABLE</th>
<th>DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local Identifiers</strong></td>
<td>Unique Student ID number</td>
<td>A stable, unique identifier for each student that allows data for an individual student to be linked across all files and submissions *See the section below regarding the Study ID.</td>
</tr>
<tr>
<td><strong>Enrollment</strong></td>
<td>School Name</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grade Level</td>
<td></td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td>Gender</td>
<td>For systems that do not have mid-year grade-level changes, once per year is acceptable</td>
</tr>
<tr>
<td></td>
<td>Race</td>
<td>All possible categories, codes, and information about how classifications are applied are needed (i.e., how handle more than one race code, race/ethnicity combined, etc.)</td>
</tr>
<tr>
<td></td>
<td>Ethnicity</td>
<td>Hispanic or Latino Not Hispanic or Latino</td>
</tr>
<tr>
<td></td>
<td>HomeLang</td>
<td>As reported by families. *If coded differently please provide codebook</td>
</tr>
<tr>
<td></td>
<td>State defined economically disadvantaged status</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Designated as limited English proficient (LEP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Special education status</td>
<td></td>
</tr>
</tbody>
</table>

Teacher data should be shared for the following academic years: 2021-2022, 2022-2023, 2023-2024.
# ADMINISTRATIVE DATA

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>VARIABLE</th>
<th>DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math Course Information</td>
<td>Teacher</td>
<td>A variable with the teacher’s name or identifier for each teacher. (optional)</td>
</tr>
<tr>
<td></td>
<td>Math course name</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Math Course attendance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Math course final grade</td>
<td></td>
</tr>
<tr>
<td>Attendance* (if available)</td>
<td>Number of Absences</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of Days Enrolled</td>
<td></td>
</tr>
<tr>
<td>Math Assessment Information</td>
<td>State Assessment Scale Score - Mathematics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>State Assessment Proficiency Level Score - Mathematics (if available)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Local Benchmark - Mathematics Assessment Name</td>
<td>Indication of which assessment the student took (e.g. grade level, name)</td>
</tr>
<tr>
<td></td>
<td>Local Benchmark Mathematics Results</td>
<td>For the assessment taken, indication of achievement, reported as a level of proficiency, scaled or raw score (in order of preference).</td>
</tr>
</tbody>
</table>

# ADDITIONAL DATA FOR STUDENTS BEING COLLECTED FOR LEARNING AGENDA TO BE SHARED

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>VARIABLE</th>
<th>DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EIC Student Survey</td>
<td>EIC Student Survey Data of Beliefs, Engagement, and Experience in Mathematics</td>
<td>All raw data</td>
</tr>
<tr>
<td>CATEGORY</td>
<td>VARIABLE*</td>
<td>DETAILS</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Teaching Site</td>
<td>School Name</td>
<td></td>
</tr>
<tr>
<td>Local Identifier</td>
<td>Unique Teacher Identifier</td>
<td>A stable, unique identifier for each teacher that allows data for an individual teacher to be linked across all files and submissions with student data</td>
</tr>
<tr>
<td>Demographics</td>
<td>T-Reside</td>
<td>How long has the teacher resided in the United States.</td>
</tr>
<tr>
<td></td>
<td>Ethnicity</td>
<td>Hispanic or Latino</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not Hispanic or Latino</td>
</tr>
<tr>
<td></td>
<td>Race</td>
<td>NIRN will need all possible categories, codes, and information about how classifications are applied (how they handle more than one race code, are race and ethnicity combined, etc.)</td>
</tr>
<tr>
<td>Education</td>
<td>National Board Certified</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Years as a teacher</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does teacher speak English fluently?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does teacher speak Spanish fluently?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the teacher speak other language fluently?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If other language, specify</td>
<td></td>
</tr>
</tbody>
</table>

### ADDITIONAL DATA FOR TEACHERS BEING COLLECTED FOR LEARNING AGENDA TO BE SHARED

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>VARIABLE</th>
<th>DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EIC Teacher Survey(s)</td>
<td>Teacher survey(s)</td>
<td>Teacher survey(s) of self-efficacy in mathematics and culturally responsive teaching, as well as implementation constructs (e.g., acceptability, feasibility) All raw data</td>
</tr>
<tr>
<td>Classroom Observation</td>
<td>Teacher observations of mathematics instruction using identified observation tool</td>
<td>Met Criteria for Fidelity or Not for each individual teacher</td>
</tr>
</tbody>
</table>
The Effective Implementation Cohort (EIC) Readiness Constructs and indicators are implementation best-practices that are shown to facilitate the uptake of evidence-based practices. This document defines the following constructs in terms of observable and measurable indicators: District Teaming, LEA Executive Sponsors, Communication, Fit and Feasibility Assessment, Implementation Planning, and Measurement Planning. Feedback from participating EIC cohort members (providers and district staff) and the Bill & Melinda Gates Foundation’s National Math Advisory Council have been incorporated into the readiness constructs’ definitions and indicators.

## EIC READINESS CONSTRUCTS AND DEFINITIONS

<table>
<thead>
<tr>
<th>READINESS CONSTRUCT</th>
<th>CONSTRUCT DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Teaming</td>
<td>A representative team has been formed at the district level and is working to lead implementation and create the enabling conditions in selected schools so that educators can make full use of high-quality middle-years math curriculum.</td>
</tr>
<tr>
<td>LEA Executive Sponsor Engagement</td>
<td>The Executive Sponsor(s) champions and supports district and school staff as they engage in implementation of high-quality middle years math curriculum. An executive sponsor is an individual with the ability to influence others, authority to make decisions regarding resource allocation, institutional knowledge, time, and positive relationships with staff and stakeholders.</td>
</tr>
<tr>
<td>Communications</td>
<td>Frequent and accurate information regarding implementation planning and progress is exchanged between stakeholders and acted upon by the identified persons. Some goals of communication include sharing information, gathering feedback and input, clarifying expectations, and celebrating successes.</td>
</tr>
<tr>
<td>Assessing Fit and Feasibility</td>
<td>An assessment conducted by districts to better understand how a new or existing high quality middle years mathematics curriculum works within their existing context to support implementation planning and use.</td>
</tr>
<tr>
<td>Implementation Planning</td>
<td>Comprehensive implementation strategies are specified within a plan to ensure capacity is developed to support successful use and sustainability of a high-quality middle years mathematics curriculum.</td>
</tr>
<tr>
<td>Measurement Planning</td>
<td>Utilize multiple methods to collect and review data to: (1) inform decision-making for continuous improvement, (2) examine effectiveness, and (3) communicate with stakeholders.</td>
</tr>
</tbody>
</table>
# DISTRICT TEAMING

## Construct Definition

A representative team has been formed at the district level and is working to lead implementation and create the enabling conditions in selected schools so that educators can make full use of high-quality middle-years math curriculum.

## Indicators

District Team is representative of the diversity of the district, schools, and priority students being served and the necessary math content knowledge, leadership/decision-making authority, cross-functional perspectives from district and schools, and provider support.

District Team utilizes effective team meeting and communication processes and procedures.

Attendance at meetings is high (greater than 80% of team members) and consistent.

District Team members have sufficient time dedicated to planning for and leading implementation functions (in and outside of meetings) or fulfilling their role and responsibilities.

Positive working relationships with building/campus leadership, school/campus teams, and staff.

District Team members follow through and accomplish identified actions within needed timelines.

District Team has clear vision, mission, and shares accountability for the work.

## Optional Resources

- **District Implementation Team Handout**
- **Guiding Questions for Team Development**
- **Active Implementation Hub Implementation Teams Learning Module**
## LEA EXECUTIVE SPONSOR ENGAGEMENT

### Construct Definition

The Executive Sponsor(s) champions and supports district and school staff as they engage in implementation of high-quality middle years math curriculum. An executive sponsor is an individual with the ability to influence others, authority to make decisions regarding resource allocation, institutional knowledge, time, and positive relationships with staff and stakeholders.

### Indicators

- Lead Partner and district team cultivate the role of LEA executive sponsor(s) and school leaders as champions
- Executive sponsor(s) and Lead Partner collaborate to support an effective partnership
- Executive sponsor(s) use data and information for planning and/or problem solving
- Executive sponsor(s) use effective bi-directional communication with district leadership, team, staff, and stakeholders (e.g., board, community)
- Executive sponsor(s) ensure that priority students are centered when making decisions for prioritizing the work, equitably allocating resources, and addressing implementation barriers
- Executive sponsor(s) can speak to and answer questions about implementation of the high-quality middle years math curriculum and motivate staff for implementation (visible promotion)
- Executive sponsor(s) is an individual with decision-making authority

### Optional Resources

- **Executive Sponsor Handout**
- **Active Implementation Hub Leadership Learning Module**
# COMMUNICATION

## Construct Definition

Frequent and accurate information regarding implementation planning and progress is exchanged between stakeholders and acted upon by the identified persons. Some goals of communication include sharing information, gathering feedback and input, clarifying expectations, and celebrating successes.

## Indicators

| District Team identifies internal and external stakeholders for bi-directional communication |
| District Team plans for communication messages, methods, and frequency for different stakeholders including school leadership, teams, and staff to celebrate successes and be responsive to identified needs in planning and implementation |
| District Team uses coherent communication messages and methods/tools as outlined in their plan |
| Effective and efficient communication processes are used between Provider, District Team, and schools |
| District Team assesses effectiveness of communication using feedback from stakeholders |

## Optional Resources

- **Active Implementation Hub Linking Communications Protocol Lesson**
- **Linking Communication Protocol Handout**
ASSESSING FIT AND FEASIBILITY

Construct Definition
An assessment conducted by districts to better understand how a new or existing high quality middle years mathematics curriculum works within their existing context to support implementation planning and use.

Indicators
Key stakeholders (e.g., school staff, students, families) are involved in identification of the needed changes related to mathematics for the priority students.

Assessment considers evidence, usability, available supports for the mathematics curriculum, how well it addresses the district's needs related to mathematics for priority students, the fit with priorities and values of the district, schools, families, and community, the district's previous experience with implementing new practices or changes, and the capacity of the district to sustain and scale implementation.

Assessment identifies initiatives with potential competing resources, capacity, and identifies mitigating factors.

Indication of assessment of policies and procedures or impact analysis related to math (e.g., access to rigorous courses, etc.) for priority students.

Hexagon Tool: Fit & Feasibility Assessment

Active Implementation Hub Hexagon Tool Process Learning Module

Optional Resources
Initiative Inventory

Active Implementation Hub Initiative Inventory Learning Module
## IMPLEMENTATION PLANNING

### Construct Definition

Comprehensive implementation strategies are specified within a plan to ensure capacity is developed to support successful use and sustainability of a high-quality middle years mathematics curriculum.

### Indicators

- Provider/LEA define shared Theory of Action to improve outcomes for priority students
- Activities are specified/identified to develop staff competency (e.g., professional learning, coaching) around the use of the chosen mathematics curriculum
- Provider and LEA demonstrate shared accountability for co-creation and use of implementation strategies
- Priority students are kept at the center of implementation planning and activities (e.g., asking who benefits from decisions and who may be further marginalized and what will this team do to mitigate this gap; how will priority students be engaged)
- Data are used to identify needs, set goals and implementation benchmarks, and monitor progress
- Provider and LEA incorporate strategies that ensure equitable resource allocation
- Planning activities progress amid competing priorities

### Optional Resources

- Implementation Stages Planning Tool
- Active Implementation Hub Implementation Stages Planning Tool Learning Lesson
- Active Implementation Hub Implementation Stages Learning Module
- Active Implementation Hub Implementation Drivers Learning Module
<table>
<thead>
<tr>
<th>Construct Definition</th>
<th>Utilize multiple methods to collect and review data to: (1) inform decision-making for continuous improvement, (2) examine effectiveness, and (3) communicate with stakeholders.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicators</td>
<td>Measurement plan includes process, programmatic, and impact measures for identified implementation goals and benchmarks.</td>
</tr>
<tr>
<td></td>
<td>Multiple methods and sources are used to collect data (e.g., observation, surveys, interviews/focus groups, assessments, administrative data) to support having appropriate and relevant evidence for decision making.</td>
</tr>
<tr>
<td></td>
<td>Indicators are well defined: leading/lagging, how data will be collected, who will collect/compile necessary data, and how frequently it will be collected.</td>
</tr>
<tr>
<td></td>
<td>Process exists for team to review, problem-solve and develop an action plan using disaggregated data as it relates to priority students and share results with stakeholders.</td>
</tr>
</tbody>
</table>

**Optional Resources**

- **Active Implementation Hub Decision Support Data System Learning Lesson**
- **Active Implementation Hub Fidelity Learning Module**