Implementation Development Map: Literature Review for Data-Driven Decision Making

October 2021

Megan Hague Angus, Rebecca Kleinman, Jessica F. Harding

Submitted to:

Bill & Melinda Gates Foundation
P.O. Box 23350
Seattle, WA 98102
Phone: (206) 709-3100
Program Officer: Sarah Weber

Submitted by:

Mathematica
P.O. Box 2393
Princeton, NJ 08543-2393
Phone: (609) 799-3535
Project Director: Jessica F. Harding

This page has been left blank for double-sided copying.

Literature Review for Data-Driven Decision Making (DDDM)

In this document, we summarize our literature review on the Data-Driven Decision Making (DDDM) element of the Implementation Development Map (IDM). We start with an overview, then provide a bulleted list that summarizes the strength of support from professional/expert recommendations and the research literature, and we discuss whether the research speaks to equity. Following the detailed notes are two graphics that summarize, for each IDM indicator, the strength of (1) the research evidence and (2) the support from expert recommendations and professional best practices. The appendix describes our literature search and review process.

A. Overview

Data-Driven Decision Making (DDDM) is one of seven elements in the Implementation Development Map (IDM). DDDM focuses on state policies around collecting and using data to inform decision making. It specifies that state data systems should allow for data linkages and centralization across programs, and that states should require programs to collect data on program quality and on a variety of student-, teacher-, classroom- and program-level measures that could be used to inform decision making and improve programs.

DDDM touts the establishment of a culture of data use to inform policies and decisions on access to and equity and availability of high quality early learning. There are nine indicators in the DDDM element and, unlike other elements in the IDM, all nine are at the state level. Indicators focus on state policy, supports, and data systems, such as statewide standards, processes for collecting data, and using data for continuous improvement.

Definitions

Research Strength is based on the number of *high quality studies* with favorable effects on child or teacher outcomes.

- High quality studies are those in which the design is strong enough to suggest that outcomes can be attributed to the intervention, practice, or policy that is being studied.

Practice Strength is based on whether the indicator is supported by professional best *practices or expert* recommendations.

- Professional best practice standards include the Head Start Performance Program Standards (HSPPS) and the standards set forth by the National Association for the Education of Young Children (NAEYC).
- Expert recommendations are from the National Academy of Sciences, Engineering, and Mathematics (NASEM).[1]

At the request of the Bill & Melinda Gates Foundation, Mathematica conducted a systematic literature review focused on DDDM. (The full methodology appears in the appendix.) For the DDDM element, after screening the studies collected for the literature review, we identified and reviewed three studies published since 2001 to assess their quality and key findings (see References). For this element, two studies supported at least one IDM indicator. Limited research examines whether access to and use of data or having a state data system are associated with classroom quality or child outcomes.

[1] Because many of the professional recommendations that we used for the other IDM elements—such as HSPPS and NAEYC standards—do not speak to data driven decision making, we have extended the documents and resources we assessed to include resources that explicitly address data driven decision making and the role it plays in building state systems to support early childhood development. As a result, we added an additional resource to support expert recommendations. The document produced by the Early Childhood Data Collaborative is a joint effort of the Center for the Study of Child Care Employment at UC Berkeley, Council of Chief State School Officers, Data Quality Campaign, the National Center for Children in Poverty at Columbia University's Mailman School of Public Health, the National Conference of State Legislatures, the National Governors Association, and the Pew Center on the States.

Despite the limited availability of high quality research, we caution readers against drawing conclusions about the inherent value of an IDM indicator. Readers should not conclude that a lack of high quality studies means that the indicator does not have valuable, nuanced information to offer about how to strengthen state systems.

Because the IDM is a tool designed to improve state systems, we also determined which elements and indicators were supported by professional best practice standards and expert recommendations. (See the box on the first page for definitions of quality, best practice standards, and expert recommendations; see the appendix for full definitions and a description of how we rated these dimensions to determine the overall research strength and practice strength of each IDM indicator.) Our assessment of the associated literature and the correspondent professional recommendations were also modified to reflect the spirit of this IDM element. Unlike other elements that referred to expert recommendations made by NASEM, our assessment of practice strength for each indicator drew on documents produced by the Early Childhood Data Collaborative.

The IDM tool explicitly embeds equity into the indicators to ensure that state leaders continue to value diverse groups of learners and provide high quality learning opportunities for all children. In our literature review, we examined equity by describing and placing value on studies that include teachers and students with diverse characteristics. We have captured whether the samples in high quality studies with favorable effects include dual language learners (DLLs), children whose families have low incomes, and children and teachers of racially and ethnically diverse backgrounds. Research that explicitly addresses questions of equity is limited, however, despite its importance for state systems that serve children from disadvantaged backgrounds.

B. Details of support for indicators

In this section, we describe the strength of support the indicators have from the research literature or the recommendations of professionals and experts. We detail the high quality studies with favorable effects, the parts of the indicator supported by the study, and any themes in the results that concern outcomes of children or teachers. We report whether any studies are particularly relevant to a specific IDM indicator and whether the research addresses equity, particularly whether studies were based on diverse samples or showed effects for certain groups of children or teachers. If there are no studies related to an indicator (Figures 2 and 4), we do not discuss it.

IDM DDDM 1: State requires programs to conduct program-level assessments using reliable and valid measurements to inform program-level continuous quality improvement (CQI). Assessments include evaluating the quality of the following six conditions:

- **Supports for dual language learners and inclusion and individualizing for children with developmental delays and disabilities**
- **Classroom environments**
- **Teacher-child interactions**
- **Curriculum implementation**
- **Family engagement practices**
- **Child outcomes (e.g., kindergarten readiness)**

Practice strength:

- Both professional and expert recommendations support programs using data to inform continuous quality improvement. HSPPS require programs to use data to identify risks and strategies for program quality improvement. Programs must develop and then implement a process to use data to identify both their strengths and needs, then evaluate their progress toward goals. Experts assert that states should build and use data to enable programs to focus on improvement.

IDM DDDM 2: State engages in DDDM to ensure high quality teaching, equitable access for children and families, equitable PD for early childhood educators, and positive child outcomes. The state's DDDM efforts include these six conditions:

- **Supporting programs to set annual (or more frequent) goals toward improving teaching and learning, equitable access, and child outcomes**
- **Monitoring programs' progress toward those goals by collecting multiple types of data, including student data (e.g., enrollment, attendance, assessments), classroom observations of teaching quality; and early childhood educators, leaders, coaches, and family surveys**
- **Disaggregating and analyzing data by targeted populations**
- **Using data for improvement of policies, and supports (e.g., funding, PD, training, etc.) to programs**
- **Supporting program leaders, early childhood educators, and other stakeholders to analyze their own data and create or modify their professional learning goals and action plans**
- **Improving data collection and data analysis processes**

Practice strength:

- Professional recommendations support this indicator. HSPSS require programs to set annual goals and evaluate their progress toward them, and NAEYC notes that an annual evaluation is an opportunity for programs to recognize strengths and identify areas for growth. HSPSS requires that programs aggregate, analyze, and compare data to help them in identifying risks and strategies to improve the program in all service areas. HSPSS also requires that child-level data be analyzed by subgroups. Expert recommendations partially support this; NASEM guidelines suggest that programs use data for evaluation and assessment but do not speak to all six of the conditions. They recommend frequent ongoing assessment and classroom observations to provide information on children's learning and development and to tailor instruction to best meet the specific needs of children, but they do not say how often this should happen or specify that programs should set annual goals.

IDM DDDM 3: State has a standardized quality rating and improvement system (QRIS) to assess program quality, or the system meets the following four conditions:

- **System includes on-site program quality assessments at least once every two or three years**
- **State system is differentiated so that programs rated lower in quality or with previous policy violations receive more frequent on-site program quality assessment visits**
- **On-site visits include classroom observations by trained and reliable observers. Observers use research-based, valid, and reliable tools to measure quality. Observations include a focus on teacher-child interactions and instructional quality.**

- **The QRIS is inclusive and aligned across multiple early learning systems, including state pre-K, private, and other early learning programs (e.g., family child care programs, child care centers). Private programs must be rated at a high level to have state pre-K classrooms or slots.**

Research strength:

- Two high quality studies with favorable effects support the use of a QRIS to rate program quality. Both studies examined the relationship between a state's QRIS ratings and program quality. One of the studies, which used a rigorous design and longitudinal data, revealed that having a lower initial rating of overall program quality in the QRIS led programs to improve their rating. This study provides causal evidence that having a QRIS with program quality ratings can lead to desired program changes. The second study found that among preschool programs participating in the QRIS, those with the highest QRIS ratings also had higher classroom quality, which suggests that QRIS program quality ratings are related to classroom quality.

Practice strength:

- The professional recommendations and the experts partially support the use of QRISs to rate program quality. HSPPS note that all programs are asked to comply with their state's QRIS requirements. Given state variation, HSPPS did not specify the four QRIS conditions listed by the indicator but rather asserted the value of Head Start programs participating in their state's QRIS system. NASEM notes that accountability systems like QRIS are a venue through which programs can improve instructional practices, although NASEM did not specifically recommend that states have QRISs.

IDM DDDM 4: State ensures access to various kinds of data on all six of the following areas:

- **Learning and development assessment data, student attendance data, including information on suspensions/expulsions**
- **Data on the qualifications and diversity of the ECE workforce**
- **Data on professional development for ECE providers, including job-embedded professional learning (JEPL) data**
- **Classroom quality data**
- **Curriculum fidelity data**
- **Data on family engagement efforts and program staff interaction/collaboration with parents**

Practice strength:

- Professional recommendations partially support this indicator. HSPPS require programs to enter many different aspects of program data into their centralized Program Information Report (PIR) system. These requirements include reporting attendance data and data on the qualifications and diversity of their workforce. PIR also requires programs to identify services offered to and received by families. The use of the Classroom Assessment Scoring System (CLASS) to monitor classroom quality is also required. However, additional data mentioned in the IDM, such as curriculum fidelity data, are not required in the PIR. NAEYC recommends that programs collect data, but the recommendations do not address state-level collection of data. Experts recommend that states collect child-level data as well as program-level data about program quality and the work environment.

IDM DDDM 5: State has formal processes for determining the relevancy and quality (i.e., reliable and valid) of data collected at the student or classroom level. These processes have been applied to all of the state's current data and data are being aggregated to the state level to use for CQI.

Practice strength:

- Both professional and expert recommendations support this indicator. As with Indicators 1 and 4, HSPPS require programs to collect and report data. Additionally, as noted, all Head Start programs are to use the data for CQI processes. Head Start also requires programs to use the CLASS, a valid assessment of classroom quality. This indicator also suggests that the state has a formal process to determine the relevancy and quality of data. For Head Start programs, the tool is provided to assess at the classroom level. NASEM notes that stakeholders should review and improve their current policies and systems for evaluation and assessment of care and improve the extent to which current evaluation and assessment procedures- including the assessment and observation tools,- assess not only children's progress but account for setting-level and community-level factors, and embedded in the continuous system of supports to inform and improve professional practice.

IDM DDDM 6: To make informed decisions, state has the infrastructure and data analytic capacity to connect different types of data to capture a full picture of the pre-K system and meets all of the following three conditions:

- **Links student data to specific classrooms**
- **System has the ability to put program data in the context of community data (e.g., demographics, family characteristics, and health)**
- **System connects professional learning data with teaching quality and child assessment data**

Practice strength:

- The professional recommendations do not support this indicator. The expert recommendations do support the indicator; they suggest that the ability to link child-level data with data in other educational and key data systems will allow policymakers to track progress over time. It will also help them deepen their understanding of the relationship between early child education programs and other child learning and development supports. Experts also recommend that states have the ability to link individual child-level data with data on the ECE workforce.

IDM DDDM 7: State has a centralized data aggregation, linking, and management system. Data management system meets the following four conditions:

- **State collects data at all appropriate levels, including classroom, program, district, and state level.**
- **State can link information across programs to account for all children served across various funding streams (e.g., child care subsidy, Head Start, Section 619-IDEA)**
- **State data system collects specific demographic data, including race, ethnicity, geography, socio-economic status, DLL status, and special needs status.**
- **State system collects and tracks longitudinal data on students to determine efficacy of pre-K efforts and collaborates with K–12 system to ensure common usable data.**

Practice strength:

- The professional recommendations partially support this indicator in that the Head Start program collects and has access to different types of data. The PIR, along with the three-year designation renewal system, which assesses all programs and ensures they meet standards across seven conditions, ensures all program-level data are stored in centralized systems. However, this reporting process and system, while centralized, is not at the state level; and the PIR system is linked only to one funding stream as opposed to several, and does not collaborate with the K–12 system to ensure common usable data. The experts recommend collecting program site data, program quality data, and work environment data, and say that data should link early childhood providers and programs.

IDM DDDM 8. State conducts regular analysis and reporting for data collected, and the system meets the following four conditions:

- **Student data are analyzed by critical subgroups (e.g., race, ethnicity, income, DLL, and children with developmental delays and disabilities)**
- **System includes the analysis of trends in the data over time and relationships between key variables. Data are used by leaders to inform decision making about policies, funding, and other supports.**
- **Collaboration with key stakeholders to interpret data, identify key issues, and gain input on plans for improvement.**
- **Identification of districts, programs, or/schools that have improved and processes for others to learn from their success.**

Practice strength:

- The professional recommendations partially support this indicator. HSPPS require programs to collect and report data. Additionally, as noted, all Head Start programs are to use the data for quality improvement processes. Although Head Start does not use data at the state level, the Head Start PIR and monitoring systems collect information across all programs. That said, although the value of including stakeholders in data use is part of the continuous improvement process, the HSPSS do not ask programs to explicitly identify other programs that have been successful in improvement efforts and processes. NAEYC recommends that programs collect data, but it does not address state-level collection of data. The experts partially support this indicator. NASEM supports state-level use of data, suggesting that states should move from using data at a point in time to tracking information over time. NASEM suggests leveraging stakeholders to help programs understand and use data and formulate improvement plans. However, identifying other programs in other states to learn from was not one of the recommendations.

IDM DDDM 9: State teams have access to data on all populations (children, early childhood educators, and parents) that are part of the pre-K system. State keeps track of equity differences in quality and achievement across identified subgroups and takes steps to eliminate those differences by disaggregating data in meaningful ways, such as by race, income, language or other important traits that historically predict inequalities, in outcome. The state's efforts to understand and address inequity include these five components:

- **Ongoing data collection,**

- **Engaging in active discussions that surface issues of inequity for targeted populations,**
- **Action planning and implementation—creating action plans and following through with implementation,**
- **Planning to assess and refine implementation**
- **Amending policies and practices that address these issues**

Practice strength:

- The professional and expert recommendations partially support this recommendation. HSPPS encourage programs to examine data as they relate to racial and ethnic groups. However, it is not clear from the standards whether programs are expected to lead active discussions, create action plans, and then implement steps to address inequity. Similarly, expert recommendations suggest using data to drive decision making to improve quality, but do not specify that programs and states are to implement such a plan to address inequities.

C. Overall ratings of research and practice support for indicators

Figures 1 and 2 summarize the overall strength of the research and practice support for each DDDM indicator.

Figure 1. Indicator key for overall ratings of reseach and practice strength

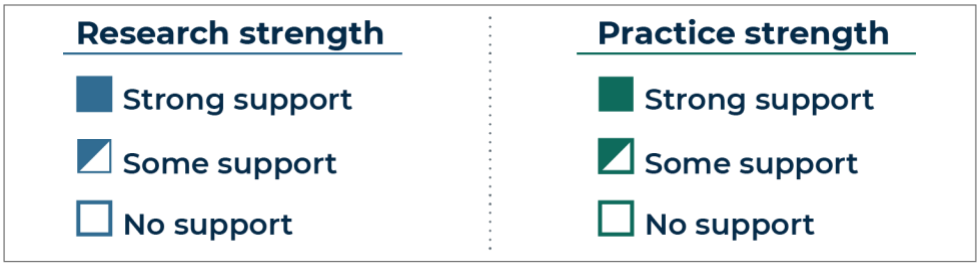


Figure 2. Overall ratings of reseach and practice strength

D. Detailed ratings of research and practice support for indicators

Figures 3 and 4 give additional detail on the research and practice support for each IDM indicator.

Figure 3. Indicator key for detailed ratings of research and practice strength

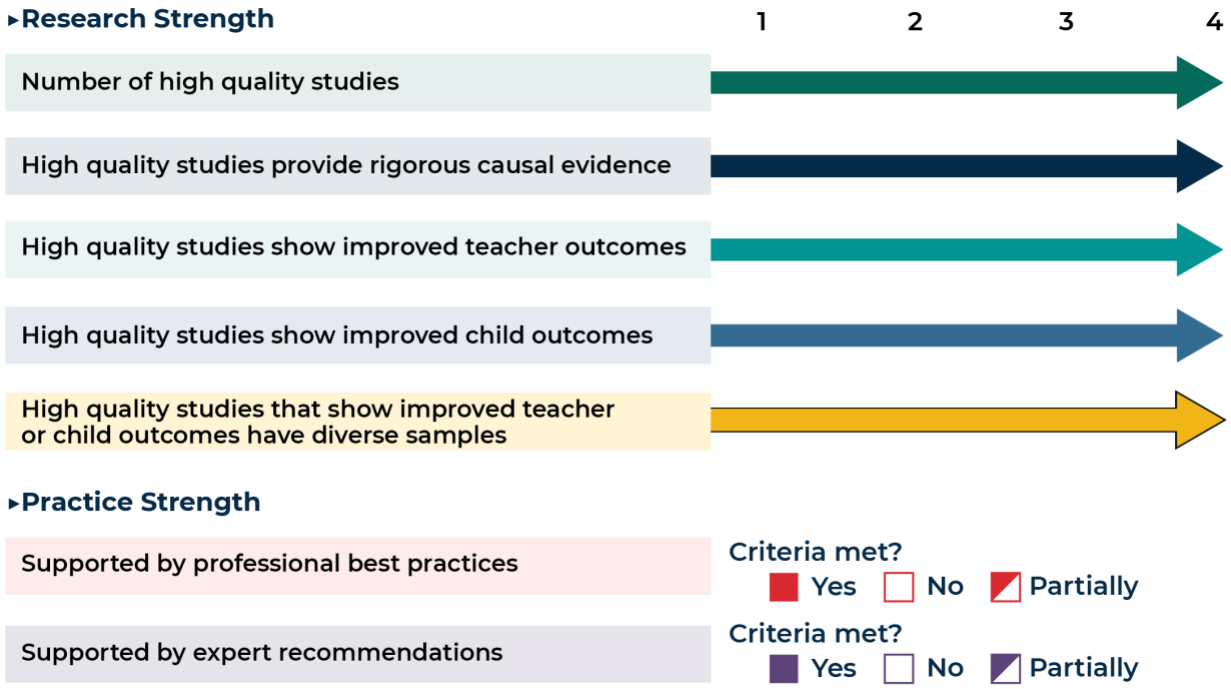
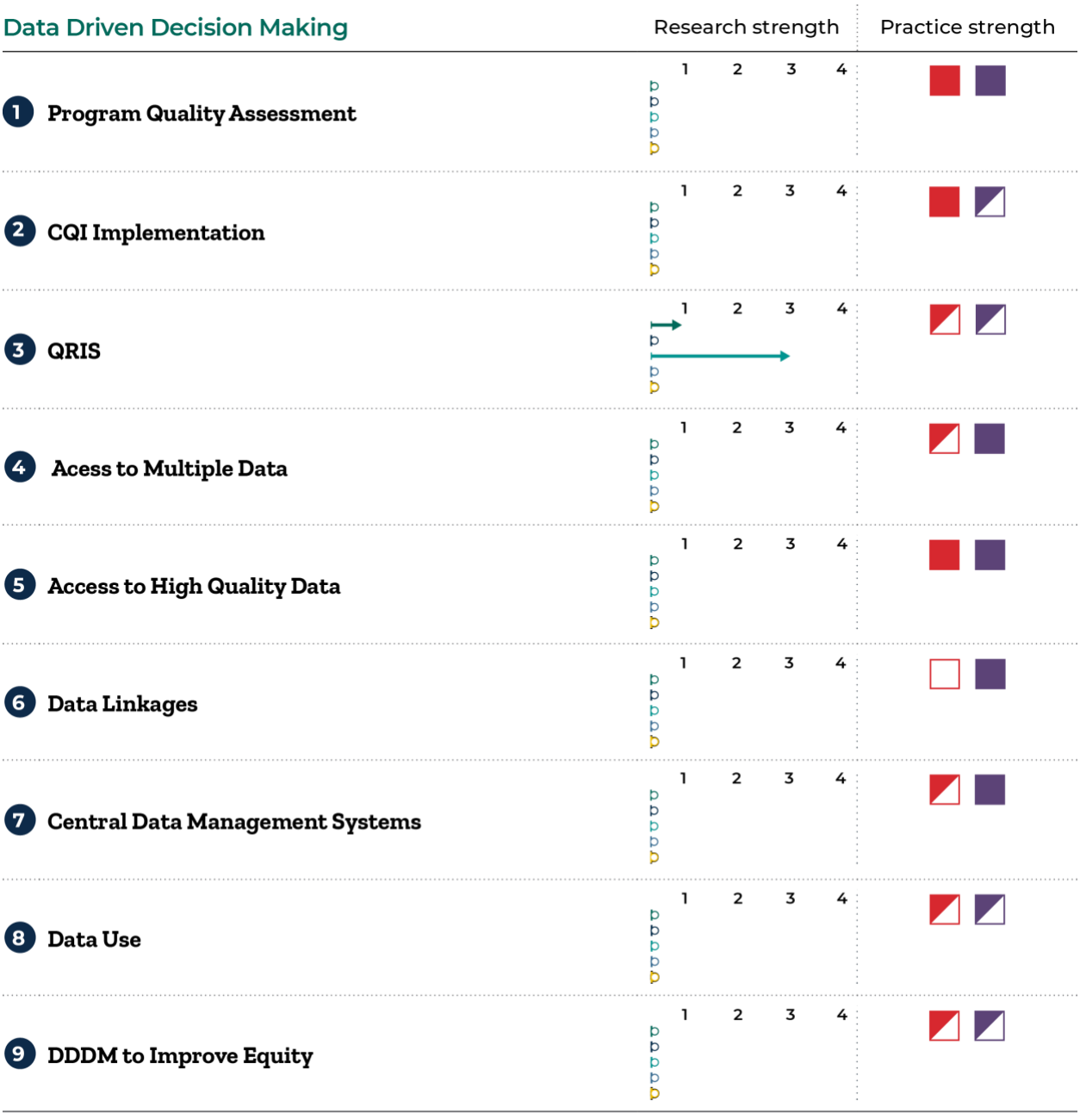


Figure 4. Detailed ratings of reseach and practice strength



Appendix

A. Identifying literature

Mathematica staff reviewed the literature on the use of research-based curriculum in preschool classrooms. We worked with our professional librarians to develop targeted search terms. We then searched eight databases for published articles.[2] Using the information in the abstracts, we screened out studies that did not meet our inclusion criteria. All eligible studies had to meet the following criteria:

- Based in the United States
- Focused on children ages 3 to 5
- Implemented in a prekindergarten setting (Head Start, child care center, or state prekindergarten program)
- Evaluated child or teacher/classroom outcomes using a randomized controlled trial, quasi-experimental, or correlational design
- Published in 2001 or later

We procured the full text of the eligible studies. Next, we screened the studies again to identify whether the studies mapped to any of the Implementation Development Map (IDM) indicators and to confirm that the studies met our inclusion criteria. We screened out any studies that did not focus on an IDM indicator (Table A.1). For the DDDM element, after examining the full texts of the nine studies initially identified, three met the inclusion criteria, all three were rated high quality, and two of the high quality studies had at least one favorable outcome (see the reference list for the high quality studies).

Table A.1. Number of studies identified and reviewed for each IDM element

IDM element	Studies identified	Studies fully reviewed	High quality studies	High quality studies with favorable outcomes
Data-Driven Decision Making	9	3	3	2

B. Assessing support for IDM indicators

We assessed each indicator on seven dimensions (Tables A.4 and A.5) to summarize the support for the indicator in the research and professional/expert recommendations.

To identify high quality studies, reviewers rated the rigor of the study design (Dimensions 1 and 2). To identify whether the studies show an improvement in outcomes, reviewers summarized the study impacts on children or teachers (Dimensions 3 and 4). To identify the extent to which high quality studies provided evidence of improvements with diverse groups of children and teachers, reviewers examined the groups of children and teachers included in the studies (Dimension 5). To determine the extent to which professional best practices and expert recommendations supported the indicators, we reviewed key practice documents (Dimensions 6 and 7). Below, we describe each step.

[2] The eight databases are Academic Search Premier, APA PsycInfo, Cochrane Database of Systematic Reviews, Education Research Complete, ERIC, ProQuest Dissertations, SAGE Journals, and Scopus.

1. Rating study quality

We wanted to identify studies with results we could be confident were valid. We categorized studies as those that provide rigorous causal evidence, strong evidence, or low quality evidence (Table A.2).

Table A.2. Study quality ratings

Study rating	Description
Provides rigorous causal evidence[a]	Well-conducted randomized controlled trials with limited attrition (< 20 percent) and no other design concerns provide the strongest evidence because outcomes can be attributed to the intervention, practice, or policy rather than to existing differences between groups.
Provides strong evidence[a]	Studies that show that their comparison groups are similar or include relevant control variables suggest that outcomes can be attributed to the intervention, practice, or policy but that unmeasured differences might exist between groups. These studies could include randomized controlled trials with high attrition or quasi-experimental designs that (a) show that the comparison groups used in analysis were equivalent on demographics and a baseline measure of the outcome (or another outcome in the same domain) or (b) controls for demographics and baseline measures. These studies could also include correlational designs and ones that have a comparison group but no baseline measures, provided they use a strong set of relevant controls (including demographics and other characteristics that could influence the outcome).
Provides low quality evidence	These are studies with unconvincing results. These studies could include randomized controlled trials with high attrition, quasi-experimental designs, or correlational studies that do not use adequate control variables or that have a confound such as using different data collection methods in the treatment and comparison groups.

[a] Both of these ratings were considered to provide high quality evidence.

We then summarized the number of high quality studies—studies that provide rigorous causal evidence and strong evidence—and the percentage of high quality studies that provide rigorous causal evidence for each indicator. Studies can support several indicators.

2. Rating study findings

We categorized whether the high quality studies had statistically significant effects on any teacher/classroom or child outcomes included in the studies (Table A.3).

Table A.3. Definitions of study impacts

Study impacts	Definition
Favorable	Significant effects on at least one outcome that benefits children or teachers/classrooms; for example, improving classroom quality
Unfavorable	Significant negative effects on at least one outcome for children or teachers/classrooms and no favorable effects on any outcomes; for example, children's receptive vocabulary scores decrease
No effect	No significant effects on any child or teacher/classroom outcomes
Mixed	At least one favorable and unfavorable effect

We next summarized for each indicator the percentage of high quality studies with favorable effects on children, teachers/classrooms, or both.

3. Rating whether studies include diverse samples

For high quality studies with favorable effects on children and teachers/classrooms, we examined whether the studies included different population groups. We assessed whether studies reported that they included the following:

- Racially/ethnically diverse children (at least 25 percent of children are Hispanic, African American, or American Indian/Alaska Native)
- Racially/ethnically diverse teachers (at least 25 percent of teachers are Hispanic, African American, or American Indian/Alaska Native)
- Children who are dual language learners (DLLs) (at least 25 percent of children are DLLs)
- Children from low-income households (at least 75 percent of children are in low-income households or the educational setting is low income)

We then looked at whether each indicator has high quality studies with favorable effects with racially/ethnically diverse children, racially/ethnically diverse teachers, DLLs, and children from low-income households.

4. Assessing professional best practices and expert recommendations

Because the IDM is a tool designed to improve state systems, we determined which elements and indicators were supported by professional best practice standards, including the Head Start Performance Program Standards, the standards set by the National Association for the Education of Young Children, and expert recommendations from the National Academy of Sciences, Engineering, and Mathematics. The latter organization analyzes available evidence to advance the learning and development of children, youth, and families and presents consensus recommendations that undergo peer review before publication.[3]

A team of researchers reviewed IDM indicators to determine how well they aligned or agreed with these professional standards. We assessed whether each indicator was supported by professional recommendations and expert recommendations by using a three-part scale that included "met," "partially met," or "not met." We used "partially met" when aspects of the indicator were supported, but not necessarily when the full indicator was met, because each indicator often covers several ideas.

5. Assigning overall ratings on dimensions

Based on the rating of study quality, study findings, the diversity of samples, and professional and expert recommendations, we rated each indicator on seven dimensions (Tables A.4 and A.5). Ratings for the

[3] See, for example: (1) Head Start Program Performance Standards (https://eclkc.ohs.acf.hhs.gov/policy/45-cfr-chap-xiii/1302-102-achieving-program-goals); (2) NAEYC, Professional standards and competencies for early childhood educators: A position statement held on behalf of the early childhood education profession, November 2019 (https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/standards_and_competencies_ps.pdf); (3) NAEYC, Early learning program accreditation standards and assessment items (https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/accreditation/early-learning/standards_assessment_2019.pdf); and (4) the Early Childhood Collaborative, Coordinated state early care and education data systems: What is next in the states?" October 2010 (https://www.ncsl.org/documents/cyf/coordinated_ece_datasystems.pdf).

research support dimensions ranged from 1 to 4; ratings for the recommendation support dimensions included met, partially met, and not met.

Table A.4. Definitions of dimension ratings

Research support dimension	1	2	3	4
Number of high quality studies	1 to 3 high quality studies	4 to 6 high quality studies	7 to 9 high quality studies	10 or more high quality studies
High quality studies that provide rigorous causal evidence	1–25% of high quality studies provide causal evidence	26–50% of high quality studies provide causal evidence	51–75% of high quality studies provide causal evidence	76–100% of high quality studies provide causal evidence
High quality studies that show improved teacher/classroom outcomes (show at least one favorable effect on a teacher outcome and no unfavorable effects)	1–25% of high quality studies show improved teacher/classroom outcomes	26–50% of high quality studies show improved teacher/classroom outcomes	51–75% of high quality studies show improved teacher/classroom outcomes	76–100% of high quality studies show improved teacher/classroom outcomes
High quality studies that show improved child outcomes (show at least one favorable effect on a child outcome and no unfavorable effects)	1–25% of high quality studies show improved child outcomes	26–50% of high quality studies show improved child outcomes	51–75% of high quality studies show improved child outcomes	76–100% of high quality studies show improved child outcomes
High quality studies that show improved teacher/classroom or child outcomes with diverse samples	Studies include one of the following groups: racially/ethnically diverse children, racially/ethnically diverse teachers, DLLs, children from low-income households	Studies include two of the following groups: racially/ethnically diverse children, racially/ethnically diverse teachers, DLLs, children from low-income households	Studies include three of the following groups: racially/ethnically diverse children, racially/ethnically diverse teachers, DLLs, children from low-income households	Studies include four of the following groups: racially/ethnically diverse children, racially/ethnically diverse teachers, DLLs, children from low-income households

DLLs = dual language learners.

Table A.5. Definitions of dimension ratings for practice support

Practice support dimension	Not met	Partially met	Met
Supported by professional best practices	The indicator was not supported by the HSPPS or NAEYC	Part of the indicator was supported by the HSPPS or NAEYC	The full indicator was supported by the HSPPS or NAEYC
Supported by expert recommendations	The indicator was not supported by NASEM	Part of the indicator was supported by NASEM	The full indicator was supported by NASEM

NAEYC = National Association for the Education of Young Children; NASEM = National Academies of Sciences, Engineering, and Medicine; HSPPS = Head Start Program Performance Standards.

6. Assigning overall ratings on research and practice strength

To make the recommendation support rating even more accessible, we summarized two dimensions of support: research strength and practice strength (Table A.6).

Table A.6. Definitions of research and practice strength ratings

Recommendation support dimension	No support	Some support	Full support
Research strength (number of high quality studies with favorable effects on child or teacher/classroom outcomes)	No high quality studies show improved child or teacher/classroom outcomes	One or two high quality studies show improved child or teacher/classroom outcomes	Three or more high quality studies show improved child or teacher/classroom outcomes
Practice strength (whether supported by professional best practices or expert recommendations)	Neither professional best practices nor expert recommendations support the indicator	At least one set of professional best practices or expert recommendations partially supports the indicator, or only one (and not both) set fully supports the indicator	Both professional best practices AND expert recommendations support the indicator

References

Bassok, D., Dee, T. S., & Latham, S. (2019). The effects of accountability incentives in early childhood education. *Journal of Policy Analysis & Management, 38*(4), 838–866.

Jeon, L., Buettner, C. K., & Hur, E. (2014). Examining pre-school classroom quality in a statewide quality rating and improvement system. *Child & Youth Care Forum, 43*(4), 469–487.

Sabol, T. J., Hong, S. L. S., Pianta, R. C., & Burchinal, M. R. (2013). Can rating pre-K programs predict children's learning? *Science, 341*(6148), 845–846.