Implementation Development Map: Literature Review for Instructional Leadership

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 Instructional Leadership (IL)

In this document, we summarize our literature review on the Instructional Leadership (IL) 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 or 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

Instructional Leadership (IL) is one of seven elements in the Implementation Development Map (IDM). IL assesses the presence of state policies, practices, and supports that enable high-quality early childhood program leadership. Instructional leaders promote high quality teaching and student learning by setting goals, allocating resources, managing curricula, monitoring lesson planning, evaluating teachers, and promoting professional development, while creating a positive, nurturing climate. There are seven indicators in the IL element and, unlike other elements in the IDM, all seven are at the state level. The IL element focuses on the actions states can take to promote high quality instructional leadership. It includes state actions to establish standards, funding, and resources for IL practices; specify leader competencies; and require programs to collect and use data on IL practices and the equitable allocation of IL resources.

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).

At the request of the Bill & Melinda Gates Foundation, Mathematica conducted a systematic literature review focused on IL. (The full methodology appears in the appendix.) For the IL element, after screening the studies collected for the literature review, we identified and reviewed 59 studies published since 2001 to assess their quality and key findings (see References). For this element, no studies supported at least one IDM indicator. This reflects the fact that the IL element represents an emerging topic in early childhood education and research on the effectiveness of emerging policies or practices is scarce. Although high quality research for some IL indicators is only limited, we caution readers against drawing conclusions about the inherent value of the indicators. The reader 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. (The box on the first page defines high 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.)

The IDM tool explicitly embeds equity into the indicators to ensure state leaders continue to value diverse groups of teachers and learners and provides 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 part(s) of the indicator supported by the study, and any themes in the results that concern outcomes of teachers and/or children. 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 teachers or children. If there are no studies related to an indicator (Figures 2 and 4), we do not discuss it.[1]

IDM IL 1: State has early childhood program standards that address the following eight instructional leadership practices:

- **Leading data-informed CQI processes**
- **Organizing and facilitating job-embedded professional learning**
- **Ensuring coherent instructional guidance and systems to support teacher practice**
- **Creating systems that support family engagement practices**
- **Ensuring effective management of operations and resources**
- **Including teachers and families in decision-making**
- **Addressing and ensuring equity for students and staff**
- **Building a trusting and supportive environment among all in the program community**

Practice strength:

- The professional and expert recommendations partially support this indicator. Both the Head Start Program Performance Standards (HSPPS) and National Association for the Education of Young Children (NAEYC) as well as the National Academies of Science, Engineering, and Medicine (NASEM) advocate for the use of early learning standards to support high quality instruction. However, these standards do not delineate the leader's role in these instructional practices, nor do they cover all the listed components.

[1] We identified one high quality study for the IL element, but it did not have favorable effects for children or teachers. It therefore does not show support for an indicator and is not detailed in this section.

IDM IL 2: Professional development is established and aligned with research-based core knowledge and competencies which align with all other applicable professional standards for early childhood instructional leaders, including credentials and degrees.

Practice strength:

- The professional and expert recommendations partially support this indicator. Both sets of recommendations note that that professional development should be offered to ECE staff and aligned with competencies and degrees however neither specifically address instructional leaders. HSPPS and NAEYC do specify which degrees or credentials are recommended for instructional leaders. NASEM partially supports this indicator as it mentions there should be competency-based qualification requirements for all ECE professionals.

IDM IL 3: State has requirements related to instructional leadership (e.g., coaching, training) in grants, contracts, regulations, or legislation for providers (e.g., program directors, site leaders) with clear guidance and/or incentives (e.g., points in a grant system, tiered reimbursement) on how requirements are to be implemented.

Practice strength:

- Professional and expert recommendations partially support this indicator. Although there is clear support among HSPPS, NAEYC and NASEM for professional development for educators, they do not offer guidance as to how providers are to offer or incentivize professional development for instructional leaders.

IDM IL 4: State requires local programs that implement instructional leadership practices to collect meaningful data (i.e., classroom observations, program evaluations, early childhood educators and family surveys), and requires the use of data collected to track progress, and to guide technical assistance and resource allocation to local programs that support improvement purposes.

Practice strength:

- Both professional and expert recommendations support this indicator. Both sets support the concept of collecting and then using data to track progress and improve programming and instruction. HSPPS, NAEYC, and NASEM all note that programs and leaders must implement a process to collect and use both classroom and program level data. Program staff and leaders can then use this data to track progress and inform CQI.

IDM IL 5: State provides specific and ongoing resources (e.g., funding for training or initiatives for pilots or innovations, technical assistance, or coaching) to support implementation of instructional leadership practices and roles. Funding and trainings are equitable (e.g., offered in diverse modes to meet the needs of the field (online, in person, accessible for staff with disabilities, Section 508 compliant).

Practice strength:

- The professional and expert recommendations partially support this indicator. HSPPS offers funding and training as well as other supports to program and center directors who might be instructional leaders through regional Office of Head Start training and technical assistance offices. NAEYC does

not speak to how best to provide resources to support instructional leadership practices and roles. NASEM supports the training and support of all educators, but it does not speak to specific supports offered to instructional leaders nor does it discuss how to deliver the training to support leaders. Neither the professional nor the expert recommendations address issues of equity in providing resources to support instructional leaders.

IDM IL 6: The state provides written guidance and resource materials to support or deliver training to support the implementation of instructional leadership practices addressing a range of topics including the following eight:

- **Leading data-informed (CQI) processes**
- **Organizing and facilitating JEPL**
- **Ensuring coherent instructional guidance and systems to support teacher practice**
- **Creating systems that support family engagement practices**
- **Ensuring effective management of operations and resources**
- **Including teachers and families in decision making**
- **Addressing and ensuring equity for students and staff**
- **Building a trusting and supportive environment among all in the program community**

Practice strength:

- Neither the professional nor the expert recommendations explicitly support the idea of written guidance or resource materials to support the training of instructional leaders in the recently mentioned topic areas. All recommendations acknowledge training should be provided to leaders and leaders should have access to training on a range of topics to promote high quality instructional environment, but they do not specify the topics, nor do they specifically say the written guidance or resource material should supplement the professional development.

IDM IL 7: State ensures that instructional leadership policies and practices promote access to high quality instructional support for all individuals. The state's efforts to understand and address inequity regarding instructional leadership include ongoing data collection, disaggregation of data, active discussions, data-driven decision-making, action planning, implementing, assessing implementation, and refining as needed. The state specifically collects data to understand and address the following three impacting factors:

- **The barriers to accessing affordable, culturally responsive, and equity-centered instructional leadership opportunities. Access includes location of available instructional leadership, language, Section 508 compliance, and diversity of trainers.**
- **Instructional leadership trainings address issues of equity, are reflective of teacher diversity, and include voices/experiences of diverse teachers and providers.**
- **Instructional leadership content is comprehensive and meets the needs of all teachers (e.g., content is varied and supports teachers to engage with children and parents from all backgrounds**).

Practice strength:

- Both the professional and expert recommendations partially support this indicator. Professional and expert recommendations assert the value of promoting access to high quality instructional support and

training to all educators. HSPPS, NAEYC, and NASEM address the value of data collection and speak to the value of collecting and using data to improve access to high quality instruction, but neither professional nor expert recommendations address disaggregating data to identify inequities or barriers to access to instructional leadership content.

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 research and practice strength

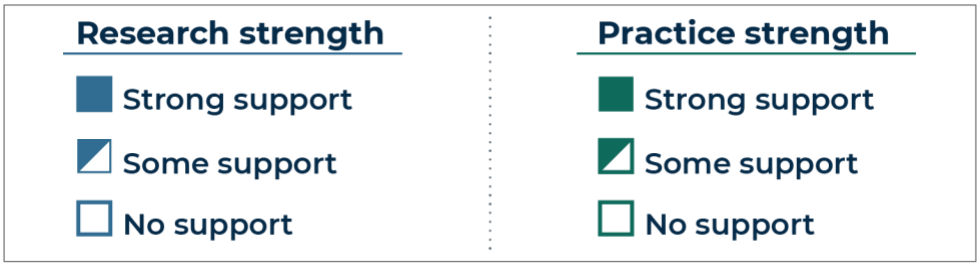


Figure 2. Overall ratings of research and practice strength

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

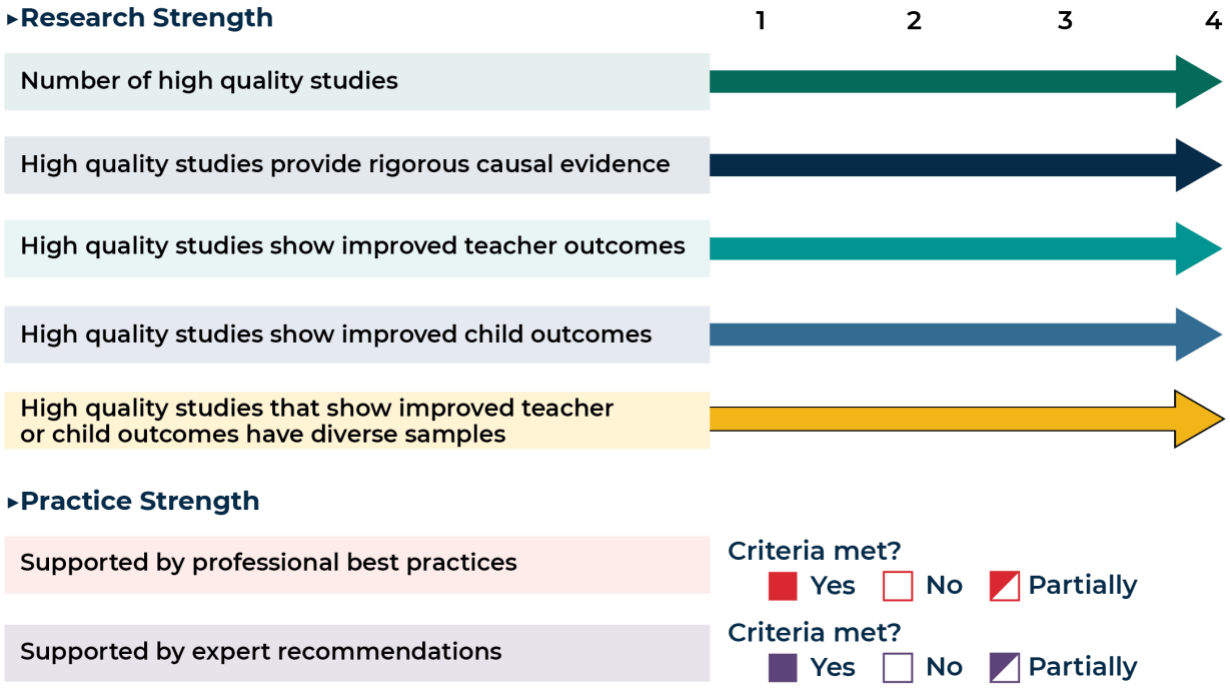
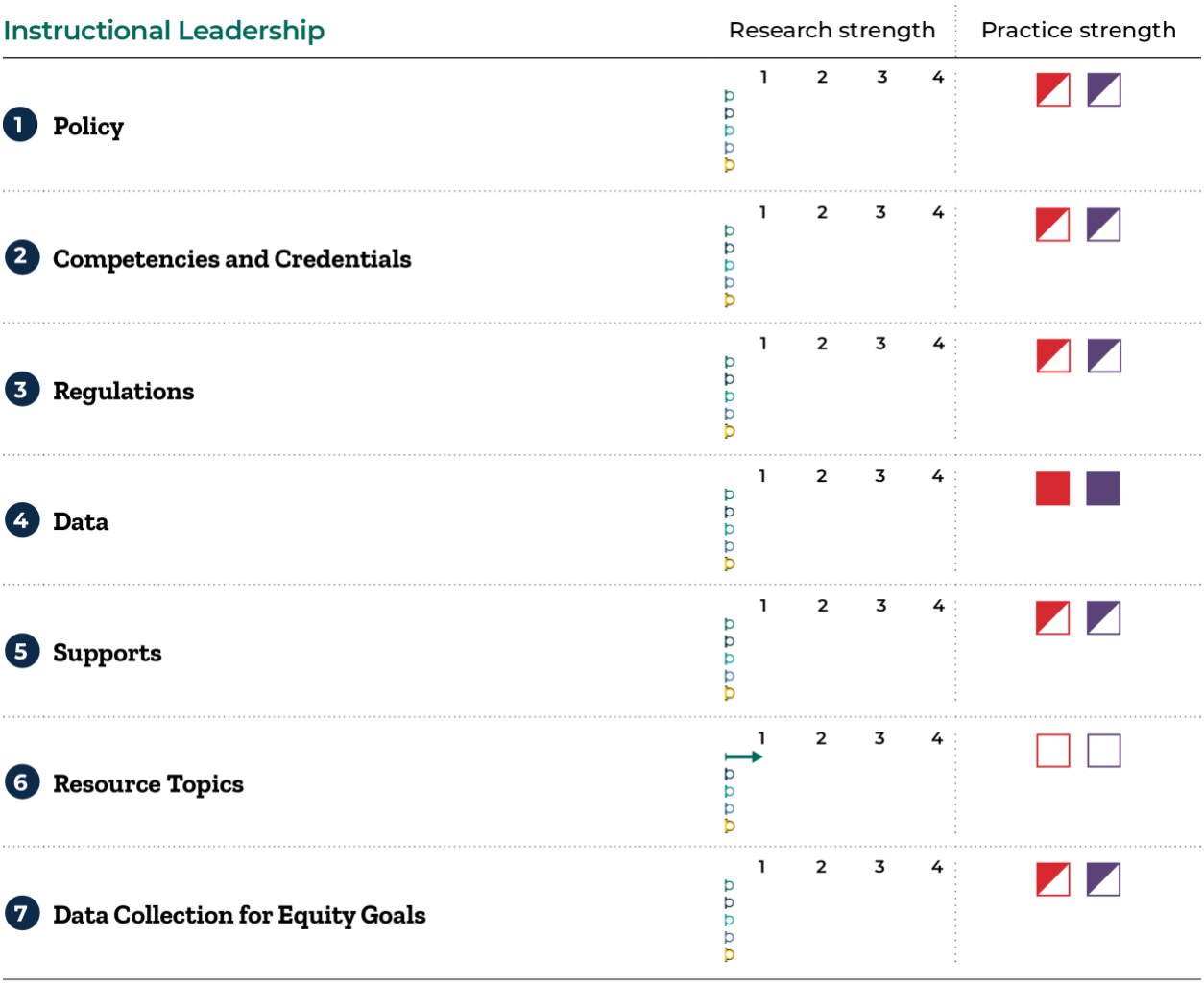


Figure 4. Detailed ratings of research and practice strength



Appendix

A. Identifying literature

Mathematica staff conducted a review of the literature focused 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 through 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 conducted a second screen 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 IL element, after examining the full text of the 35 studies initially identified, one met the inclusion criteria, and it was rated high quality; the high quality study did not have at least one favorable outcome (see the reference list for the high quality study). Thus, in Section B, we did not summarize research findings for any of the indicators.

Table A.1. Number of studies identified, reviewed, and found to support the IL element

IDM element	Studies identified	Studies fully reviewed	High quality studies	High quality studies with favorable outcomes
Instructional Leadership	35	1	1	0

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 and/or teachers (Dimension 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

[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.

professional best practices and expert recommendations supported the indicators, we reviewed key practice documents (Dimensions 6 and 7). Below, we describe each step.

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 child or teacher/classroom 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.

[3] See, for example: (1) Head Start Program Performance Standards (https://eclkc.ohs.acf.hhs.gov/policy/45-cfr-chap-xiii/1302-92-training-professional-development); (2) National Association for the Education of Young Children (NAEYC), **"**Professional Standards and Competencies for Early Childhood Educators: A Position Statement Held on Behalf of the Early Childhood Education Profession (Washington, DC: NAEYC, 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) National Research Council, "Transforming the Workforce for Children Birth Through Age 8: A Unifying Foundation" (Washington, DC: National Academies Press, 2015).

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 (Table A.4 and Table A.5). Ratings for the 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 for research support

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 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 dimensions	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

Whalen, S. P., Horsley, H. L., Parkinson, K. K., & Pacchiano, D. (2016). A development evaluation study of a professional development initiative to strengthen organizational conditions in early education settings. *Journal of Applied Research on Children, 7*(2), 1–43.