WIDELY USED MEASURES OF PRE-K CLASSROOM QUALITY

What We Know, Gaps in the Field, and Promising New Directions

By Christina Weiland and Paola Guerrero Rosada, University of Michigan

High-quality pre-K sets children on a path for success in kindergarten and beyond. Yet most children in the United States do not attend high-quality programs, with particularly pronounced gaps in quality for children from families with low incomes, dual language learners, Black children, and Latino children.[1]

Ensuring equitable access to high-quality pre-K requires being able to *measure* quality, particularly at a large scale. Accurate, reliable, and timely measurement of pre-K quality can help parents make better choices when they have multiple early learning options, and can help policymakers, school administrators, and educators improve the quality of early education settings.[2] At its core, early childhood measurement is an equity issue.

Although there are many available measures of the learning experiences of children in pre-K, only a few have been widely used. As reviewed in this brief, these commonly used measures have many strengths, such as facilitating cross-system comparisons of quality and identifying areas for teachers' growth. However, existing widely used measures tend to be weak predictors of gains in children's skills in pre-K, overlook individual learning experiences, pay limited attention to cultural responsiveness, and typically do not measure the content of children's learning. Policymakers, program administrators, and teachers can only respond to what is measured. Because of these shortcomings, when used in isolation, these measures may unintentionally hold back progress critical to closing gaps in early learning quality, equity, and opportunity.

These limitations have led experts to call for a new generation of quality measurement work in early childhood.[3] Some of this work has already begun. This brief describes the existing landscape of widely used measures of pre-K quality, further spotlights some of the newer measurement work, and concludes with a discussion of future directions for the field. Investing now to strengthen measures of pre-K quality is critical to building equitable early learning opportunities for children from historically marginalized groups.

Widely Used Measures of Pre-K Quality: Strengths and Opportunities for Improvement

Observational tools for measuring classroom quality have been used in the early childhood field for decades. The two most widely used measures are the Early Childhood Environment Rating Scale (ECERS-3 and ECERS-R) and the Classroom Assessment and Improvement System (CLASS): Pre-K.[4] Versions of these measures are available for younger and older age groups as well as for use in home-based family child care. The focus of this brief is on measures used in center-based pre-K settings.

The ECERS-3 measures the overall quality of pre-K classrooms using six scales: Space and Furnishings, Personal Care Routines, Language and Literacy, Learning Activities, Interactions, and Program Structure. Importantly, the tool's most recent form introduced items on inclusiveness and cultural sensitivity.

The CLASS Pre-K measures the quality of interactions between teachers and children and between children and their classmates in three areas, or domains: Emotional Support, Classroom Organization, and Instructional Support. These scales measure whether a classroom's emotional climate is positive, warm, and child-centered; whether time management, teacher classroom management, and classroom materials optimize children's engagement in the learning activities; and whether classroom conversations and teacher feedback and questions deepen children's thinking.

For both the ECERS and CLASS measures, data are collected by a reliable, trained assessor on a typical day in the classroom or by coding a videotape of a typical day. The length of observation varies but generally includes at least an 80-minute session. Table 1 summarizes the domains of each measure and provides information on observation requirements. Importantly, new versions of both tools are being developed but are not yet available.

Strengths of Widely Used Measures

The ECERS and CLASS tools have five main strengths:

1 THEY CREATE A COMMON LANGUAGE. Early education is characterized by fragmented systems and multiple funding streams. Common measures have helped to unite practitioners, administrators, policymakers, and researchers around a shared understanding of what high-quality pre-K is. The CLASS's Emotional Support scale, for instance, has helped elevate this important dimension of classrooms and define it as the degree to which teachers are responsive to children, the classroom climate is welcoming and warm, and teachers respect children's perspectives and experiences. Accordingly, though Head Start and state

Table 1. Comparison of ECERS-3 and CLASS

	ECERS-3		CLASS	
FOCUS	Structural and Process Quality (Classrooms ages 3-5)		Process Quality (Prekindergarten classrooms ages 3-5)	
SCALES	Space and Furnishings	Developmentally appropriate furnishing and equipment, arrangements, privacy, and child-oriented display	Emotional Support	Extent to which classroom climate is positive for learning, and teachers are sensitive and respectful of students needs and perspectives
	Personal Care Routines	Safety and health practices, interactions during routines such as meals and toileting	Classroom Organization	Ways in which teachers optimize learning opportunities and maintain children's engagement with learning through management strategies and materials
	Language and Literacy	Opportunities to expand vocabulary, use language, become familiar with print, explore and use books with staff and independently	Instructional Support	Extent to which interactions promote children's cognitive and language development through modeling, conversation, and learning activities
	Learning Activities	Access, use, and engagement with materials; opportunities for numeracy and familiarity with printed numbers		
	Teaching Interactions	Opportunities for exploration and creativity, individualized learning, practice and follow-up. Supervision and discipline strategies		
	Program Structure	Balance between transitions and waiting times, opportunities for play, and whole-group learning		
ITEMS	35 main items, scored from 1 to 7. Each item contains between 3 and 11 indicators to aid the scoring process.		1- to 7-point Likert scales with anchored-in descriptions of low-, moderate-, and high-quality classroom interactions	

Table 1 (continued)

	ECERS-3	CLASS
METHOD	Direct observation and documents review	Direct observation
DURATION	Three hours, with some additional time to examine materials and areas not observed during the observation period	Varies, but generally requires four 20-minute cycles
OBSERVER	Trained, reliable (within one point) observers	Trained observers with up-to-date reliability certification

NOTES: ECERS-3 = Early Childhood Environment Rating System, Third Edition.
   CLASS = Classroom Assessment Scoring System.

pre-K systems may differ greatly, stakeholders can share lessons about what changes and practices have helped sustain or improve scores on quality dimensions over time.

2   **THEY IMPROVE QUALITY.** When pre-K accountability systems include observational quality measures, programs and teachers do show improvements on the measures over time.[5] In intervention studies, too, when professional development is tied to these measures, pre-K teachers also show marked improvements when compared with business-as-usual practices.[6]

3   **THEY PROMOTE EQUITY AT THE SYSTEMS LEVEL.** These measures also have been used to show disparities in access to quality programs by family income, rurality, and race/ethnicity.[7] Identifying these gaps should be a priority for policymakers and practitioners. Without these existing measures of pre-K quality, there would not be comparable data across settings.

4   **THEY FACILITATE CROSS-SYSTEM COMPARABILITY.** The Institute of Education Sciences has urged researchers to use the same measures to support cross-study comparisons.[8] Using the ECERS and CLASS has already brought this advantage to early childhood education. For example, out of 45 states currently implementing a quality rating and improvement system, 19 use ECERS-3 (or ECERS-R) and 11 use CLASS.[9] Figure 1 displays this cross-systems advantage and the relative strengths and weaknesses of different city, state, and national pre-K systems using these tools.

Figure 1. Cross-System Comparisons Using ECERS and CLASS Scores
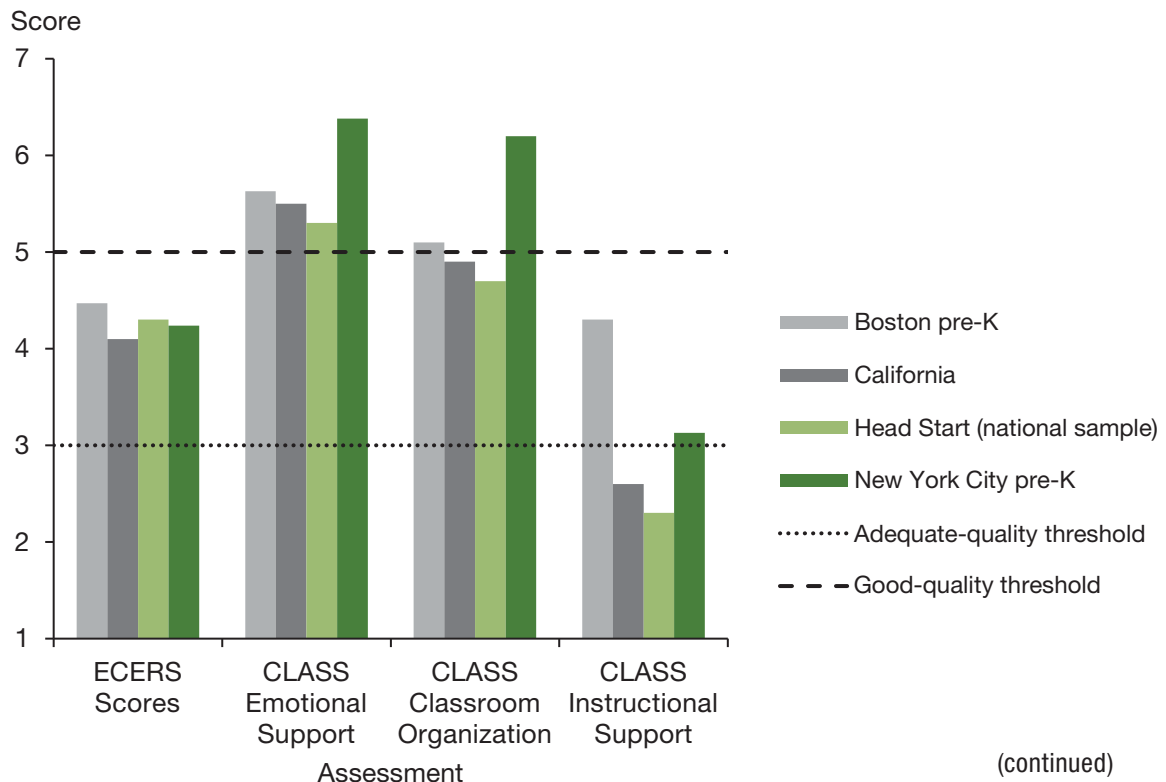
(continued)

Figure 1 (continued)

SOURCES: The data are from Joanne Hope Denny, Rena Hallam, and Karen Homer, "A Multi-Instrument Examination of Preschool Classroom Quality and the Relationship Between Program, Classroom, and Teacher Characteristics," *Early Education and Development* 23, 5 (2012): 678-696; Lynn A. Karoly, Bonnie Ghosh-Dastidar, Gail L. Zellman, Michal Perlman, and Lynda Fernyhough, *Prepared to Learn: The Nature and Quality of Early Care and Education for Preschool-Age Children in California* (Santa Monica, CA: RAND Corporation, 2008); Scott Latham, Sean P. Corcoran, Carolyn Sattin-Bajaj, and Jennifer L. Jennings, "Racial Disparities in Pre-K Quality: Evidence from New York City's Universal Pre-K Program," *Educational Researcher* 50, 9 (2021): 607–617; Emily Moiduddin, Nikki Aikens, Louisa Tarullo, Jerry West, Yange Xue, and Jerry West, *Child Outcomes and Classroom Quality in FACES 2009* (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services); Christina Weiland, Kchersti Ulvestad, Jason Sachs, and Hirokazu Yoshikawa, "Associations Between Classroom Quality and Children's Vocabulary and Executive Function Skills in an Urban Public Prekindergarten Program," *Early Childhood Research Quarterly* 28, 2 (2013): 199–209.

NOTES: The thresholds for adequate quality (3) and good quality (5) come from existing work to identify key thresholds associated with gains in child outcomes. For example, see Margaret Burchinal, Nathan Vandergrift, Robert Pianta, and Andrew Mashburn, "Threshold Analysis of Association Between Child Care Quality and Child Outcomes for Low-Income Children in Pre-Kindergarten Programs," *Early Childhood Research Quarterly* 25, 2 (2010): 166-176.
    ECERS = Early Childhood Environment Rating Scale.
    CLASS = Classroom Assessment Scoring System.

5   **THEY IDENTIFY STRENGTHS AND AREAS FOR GROWTH.** Existing pre-K quality measures can also help pinpoint what systems, programs, and teachers are doing well and what areas need further strengthening. For example, as shown in Figure 1, programs across the country are generally doing a good job on Emotional Support and Classroom Organization but not on delivering high-quality Instructional Support. These descriptive findings help to "take the temperature" of the programs and point to a particular need to examine instructional practices, particularly in pre-K.[10] Notably, the Boston pre-K program recorded the highest Instructional Support scores after it had invested intentionally in evidence-based curriculum and coaching.[11]

Limitations of Widely Used Measures

No measure is perfect. All measures have limitations, and the usefulness of each tool depends on its purpose and the question at hand. The ECERS and CLASS tools currently have several critical limitations that affect their ability to set the course for improving quality and children's learning at a large scale:

1   **THEY DON'T PREDICT GAINS IN CHILDREN'S ASSESSED SKILLS VERY WELL.** Both ECERS and CLASS are fairly weak predictors of children's learning gains in pre-K.[12] A recent study in Boston found no links between the CLASS scores and gains in language, math, and executive function across different modeling approaches and robustness checks.[13] Some studies have found slightly stronger associations between CLASS scores and outcomes in samples of native English-speaking children, children without individualized education programs (IEPs), nonimmigrant children, and children in families with higher incomes. The reasons for these differential associations are still unknown.[14] In part, the lack of predictive power may be because these measures tend to be collected on just one day, increasing the likelihood

of rater effects and other sources of measurement error.[15] Recent evidence suggests that more days of observation may be needed, but this may not be realistic or practical in terms of time, cost, and disruption to classrooms.[16] Less obtrusive approaches like cameras on the ceiling in place of human observers in the classroom may alleviate these issues, though obtaining consent from families and teachers presents hurdles.

2 **THEY DON'T MEASURE CHILDREN'S INDIVIDUAL LEARNING EXPERIENCES.** The ECERS and CLASS tools provide measures at the classroom level, averaged across all children and the teacher or teachers. However, research shows that the individual learning experiences of children in the same pre-K classrooms can vary a great deal and this variation can predict their learning gains.[17] These measures also cannot identify *how* children's learning experiences may vary systematically across student characteristics, a fact that is troubling in terms of equity. For example, studies that have examined children's individual learning experiences have found that Black boys experience more negative interactions in pre-K classrooms and are more likely than their peers to be expelled, and that girls spend more time on literacy activities than boys.[18] This is an important missing piece in improving quality by using these tools.

3 **THEY PROVIDE LIMITED INFORMATION ON CULTURAL RESPONSIVENESS AND INCLUSION.** The CLASS tool does not measure the degree to which children's cultures are reflected and focused on in classroom practice, nor the degree to which children with disabilities are fully included. (See Box 1 for definitions of cultural responsiveness and inclusion.) In contrast, ECERS *does* include indicators capturing whether materials and routines reflect classrooms' diversity in a positive way and are free of stereotypes. It also includes items measuring acceptance of diversity and whether there is evidence of bias in interactions and activities. These indicators, however, inform neither the extent to which children's culture, language, heritage, and experiences are valued and incorporated into learning activities and routines nor the

---

Box 1. Defining Cultural Responsiveness and Inclusion

Cultural Responsiveness refers to teachers' practices that connect the living experiences and perspectives of ethnically diverse students with their classroom experiences to foster their social, emotional, cognitive, linguistic, and physical development. Culturally responsive teaching builds learning communities, responds to ethnic diversity in the delivery of instruction, and reflects and models fairness and justice.

Inclusion refers not just to access to the same classroom space for children with and without disabilities but to full and equal participation, social relationships, and learning outcomes—that is, true equal access to resources and experiences beyond simple physical placement.

SOURCES: National Association for the Education of Young Children (NAEYC), "Advancing Equity in Early Childhood Education: A Position Statement of the National Association for the Education of Young Children" (Washington, DC: NAEYC, 2019); U.S. Department of Health and Human Services/Department of Education, "Policy Statement on Inclusion of Children with Disabilities in Early Childhood Programs" (Washington, DC: U.S. Department of Health and Human Services, Department of Education, 2015).

extent to which all children are affirmed and offered equitable opportunities to engage in positive interactions. This gap is especially noteworthy, given that pre-K programs are increasingly more diverse in terms of children's cultural and linguistic backgrounds and the inclusion of children with disabilities.[19] Research has shown that implementing a culturally responsive curriculum or using culturally responsive instruction can help decrease stress and anxiety that undermines learning.[20] This may support children's social and language skills and can enhance relationships between teachers and students, promote children's academic engagement, and strengthen their critical thinking.[21]

4 **THEY DON'T MEASURE CURRICULUM CONTENT, SCOPE, OR SEQUENCE.** The most effective pre-K curricula follow developmental science regarding how children develop skills specifically within a given learning domain such as language, literacy, and math.[22] They also share a focus on *content*. For example, in math, effective curricula go beyond simple counting to include deeper mathematical thinking and geometry.[23] In literacy, as shown in Box 2, curricula may go beyond teaching letter sounds in isolation to teaching them in the context of rich vocabulary. However, ECERS and CLASS do not measure these features of classroom instruction. Accordingly, the feedback they provide to teachers may not be specific or targeted enough to improve children's learning.

Why These Gaps Matter

When used in isolation, the ECERS and CLASS tools may (unintentionally) hold back progress critical to closing disparities in early learning. For example, including these measures in accountability systems has led to coaching and professional development centered on improving scores on these tools. But a randomized trial in which teachers were coached on the CLASS tool found no impacts on important child outcomes like early language and literacy skills.[24] Further, while Head Start programs have been held accountable for CLASS scores since 2012, average CLASS scores overall have remained flat since 2014, as illustrated in Figure 2.[25] Investments that aim to improve scores on these existing quality measures are unlikely to make a major dent in the large, consequential disparities in school readiness skills currently disadvantaging children from racially, socioeconomically, and linguistically marginalized backgrounds.[26]

In contrast, coaching focused on improving the implementation of a proven curriculum—a much more effective approach to improving children's learning—is largely unheard of outside of demonstration trials and in a few localities. [27] Similarly, programs are generally not incentivized to either increase their cultural responsiveness or to take a hard look at inequities in individual children's classroom experiences.

Programs and teachers respond to what is measured. Shifting programs to more evidence-based, equity-centered teaching and learning models and investments requires measures of those factors.

Box 2. Defining Cultural Responsiveness and Inclusion: A Comparison of Two Hypothetical Classrooms

Classroom A: Skills-Based Instruction →

- The lead teacher puts each letter of the alphabet on a flashcard with a picture of a word that starts with that letter. There is no theme or connection between the words (for example, "kangaroo" for "K" and "jelly" for "J").

- In whole-group circle time, the lead teacher calls on children one by one to say the letter sound that corresponds with each word on the flashcard.

This lesson helps children know their letter sounds but there is no clear learning goal beyond that.

← Classroom B: Rich, Relevant, and Engaging Instruction

- The classroom is engaged in a series of lessons based on a theme, such as the ocean as a habitat. Most lessons for a six-week period are focused on helping children develop deep knowledge on this theme.

- The teacher puts each letter of the alphabet on a flashcard with a picture of a plant or animal that lives in the ocean.

- Children are asked to say the letter sound that corresponds to the picture on the flashcard. The teacher follows up each response with a targeted vocabulary word from this particular series of lessons linking that plant or animal to understanding the ocean as a habitat.

- This activity segues into a discussion about why animals and plants have different types of habitats even when they all live in the ocean.

Children are supported in learning letter sounds—a critical early literacy skill—but are also exposed to rich content that they are engaged in outside of this lesson, new vocabulary, and supports for background knowledge and critical thinking.

Figure 2. Head Start Grantees' CLASS Scores from 2012-2018

Average Score
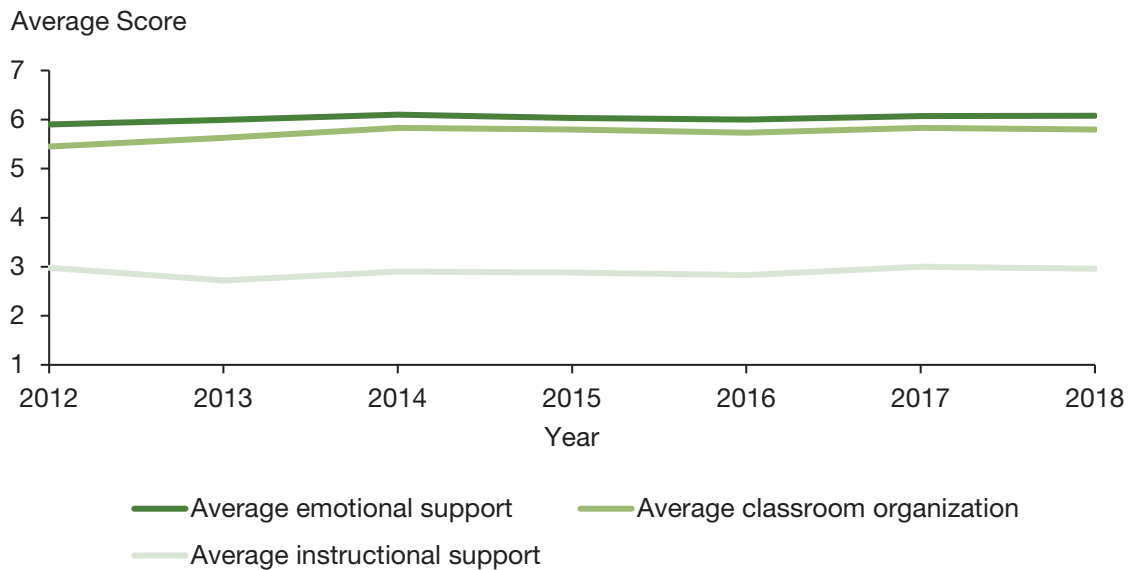
SOURCE: Office of Head Start Early Childhood Learning and Knowledge Center.

NOTE: Data points for 2019 and 2020 are excluded because there were fewer observations due to the COVID-19 pandemic.

New(er) Directions

There are many other tools for measuring children's learning experiences that address one or more of the limitations of more widely used measures. Such limitations have led researchers to call for a new generation of classroom quality measurement work, some of which has already begun. Several of these measures are summarized in Table 2, followed by a discussion of a few exemplars that address critical gaps in the ECERS and the CLASS tools.

Measuring Children's Individual Experiences: The ISI

As shown in Table 2, the most comprehensive measure of children's individual classroom experiences is the Individualizing Student Instruction Measure (ISI).[28] The ISI requires trained coders to watch videotapes of classroom instruction and then code second-by-second what individual children in the classroom are doing. The ISI measures the amount of time each child engages in math, language, and literacy learning activities, the amount of time each child is exposed to content areas, and the amount of time each child spends in different learning formats (whole group, small group, centers, and individual work). It can also be customized for an individual team or locality's needs. For example, in a Boston study, the research team worked with language and math experts to add more detailed codes of important dimensions of early education classrooms, as well as more detailed codes of teachers' organizational strategies.[29]

Table 2. Measures Capturing Children's Individual Experiences, Culturally Responsive Interactions or Inclusion, and Learning Content

MEASURE	TARGET	IE	CRI	LC	PREDICTIVE PROPERTIES	TIME-SAMPLING CRITERIA
MEASURES THAT ADDRESS TWO GAPS IN WIDELY USED MEASURES						
Dual Language Learners Discourse Snapshot (DUALLS)[a]	Children's experiences of patterns of teachers' use of discourse strategies	Yes	Yes	No	Pre-K teachers' use of "didactic and language modeling strategies" was positively related to children's bilingual expressive vocabulary skills and negatively related to children's positive engagement. Teachers' use of "elaborative and responsive language" was positively related to children's positive engagement. To date, there is no evidence of associations with children's gains in skills.	Observers rate four children per classroom in a three- to four-hour observation, alternating turns of five-minute cycles, for an average of 24 observations (six cycles per child).
Individualizing Student Instruction (ISI)[b]	Amount of time each child engages in learning activities, amount of time each child is exposed to content areas, and amount of time each child spends in different learning formats	Yes	No	Yes	Time spent on whole-group and individual activities led by a teacher predicted early literacy gains, and child-led experiences predicted vocabulary gains. However, relations with children's language, literacy, math, and executive function gains were mostly null in another study. This instrument also captures differences in experiences across pre-K students enrolled in the same classroom and across student subgroups.	Classrooms are videotaped for approximately two hours, and a coder records each target child's experiences second-by-second.

(continued)

Table 2 (continued)

MEASURE	TARGET	IE	CRI	LC	PREDICTIVE PROPERTIES	TIME-SAMPLING CRITERIA
MEASURES OF CHILDREN'S INDIVIDUAL LEARNING EXPERIENCES						
Individualized Classroom Assessment Scoring System (inCLASS)[d]	Children's competence during everyday interactions with teachers, peers, and tasks in a pre-K classroom environment	Yes	No	No	Children's positive engagement with teachers was related to gains in executive function, and children's active engagement with tasks was associated with gains in emotion regulation.	Each observation cycle consists of 10–15 minutes of watching and note-taking, followed by five minutes of scoring.
MEASURES OF CULTURALLY RESPONSIVE INTERACTIONS OR INCLUSION						
Assessing Classroom Sociocultural Equity Scale (ACSES)[e]	Equitable sociocultural interactions in early childhood classrooms with racially marginalized learners (RMLs)	No	Yes	No	Three of the five scales showed positive predictive relations with gains in children's math, executive function, and/or social competence. One scale – equitable learning opportunities – predicted slower growth in children's math and executive function. Relations between the scales and children's problem behavior were null.	ACSES is scored from 1-4 video recordings of classroom instruction, each approximately 15 minutes long. There needs to be at least one racially marginalized learner in the classroom for videos to be scored.
Inclusive Classroom Profile (ICP)[f]	Quality of classroom practices that support the developmental needs of children ages 2 to 5 with disabilities in early childhood settings.	No	Yes	No	No published evidence of predictive validity.	Classrooms are rated during a 2½- to 3-hr observation. The majority of the items are assessed through direct observation of daily classroom routines. A few items are assessed through a teacher interview and a review of documents such as a program's inclusion-related policies and tools.

(continued)

Table 2 (continued)

MEASURE	TARGET	IE	CRI	LC	PREDICTIVE PROPERTIES	TIME-SAMPLING CRITERIA
MEASURES OF CONTENT						
Boston Fidelity Tool[g]	Content-rich instruction and cognitive demand in prekindergarten classrooms	No	No	Yes	Cognitive demand and content-rich instruction predicted gains in math skills.	Live or videotaped observations of instructional time (2-3 hours) with focus on language/literacy and mathematics.
Classroom Observation of Early Mathematics Environment and Teaching (COEMET)[h]	Classroom math culture and quality of math activities	No	No	Yes	COEMET scores predicted gains in children's math skills, specifically in number recognition, geometry, and algebra (that is, patterns) tasks.	Assessors spend no less than one half-day in the classroom, aiming for a period when mathematics is taught. All math activities during that time frame are observed and rated.
Early Childhood Language and Literacy Classroom Observation Tool, Pre-K (ELLCO Pre-K)[i]	Quality of literacy environments. The ELLCO consists of three parts: literacy environment checklist, classroom observation, and literacy activities rating scale	No	No	Yes	ELLCO scores have shown associations with gains in early reading. Work on literacy areas have shown to be associated with children's gains in alphabet knowledge. Language, literacy and curriculum predicted children's gains in early reading; Activities and Environment predicted gains in Woodcock-Muñoz Language scores.	The literacy environment checklist can be completed in approximately 15 minutes. The classroom observation takes approximately 45 minutes, and 10 minutes are needed to score the literacy activities rating scale.
Observation of Language and Literacy Instruction (OLLI)[j]	Frequency, intensity, and quality of an extensive range of language and reading comprehension teaching practices	No	No	Yes	Four practices predicted children's gains in language and comprehension outcomes: engaging students in defining new words, focusing on the meaning of texts, word knowledge, and higher-order thinking.	Each observation consists of six 15-minute segments, for a total of 90 minutes of observed instruction per session, and five additional minutes per segment to score items.

(continued)

Table 2 (continued)

SOURCES: [a]Although the DUALLS does not measure culturally responsive interactions, the measure is designed to capture teacher practices specifically intended to meet the needs of dual language learners. See Natalia M. Rojas, Pamela Morris, and Amudha Balaraman, "Finding Rigor Within a Large-Scale Expansion of Preschool to Test Impacts of a Professional Development Program," *AERA Open* 6, 4 (2020).

[b]Carol McDonald Connor, Frederick J. Morrison, Barry J. Fishman, Claire Cameron Ponitz, Stephanie Glasney, Phyllis S. Underwood, Shayne B. Piasta, Elizabeth Coyne Crowe, and Christopher Schatschneider, "The ISI Classroom Observation System: Examining the Literacy Instruction Provided to Individual Students," *Educational Researcher* 38, 2 (2009): 85-99; Christina Weiland, Lillie Moffet, Paola Guerrero Rosada, Amanda Weissman, K. Zhang, Michelle Maier, Catherine Snow, Meghan McCormick, JoAnn Hsueh, and Jason Sachs, "Is Child-Level Measurement the Key to Improving the Predictive Validity of Observational Measures in Early Education Classrooms?" unpublished paper (New York: MDRC, 2021).

[c]Sally Atkins-Burnett, Susan Sprachman, Michael Lopez, Margaret Caspe, and Katie Fallin, *The Language Interaction Snapshot (LISn): A New Observational Measure for Assessing Language Interactions in Linguistically Diverse Early Childhood Programs* (Princeton, NJ: Mathematica Policy Research, 2010); Mary Bratsch-Hines, Margaret Burchinal, Ellen Peisner-Feinberg, and Ximena Franco, "Frequency of Instructional Practices in Rural Prekindergarten Classrooms and Associations with Child Language and Literacy Skills," *Early Childhood Research Quarterly* 47 (2019): 74-88.

[d]Jason T Downer, Leslie M. Booren, Olivia K. Lima, Amy E. Luckner, and Robert C. Pianta, "The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary Reliability and Validity of a System for Observing Preschoolers' Competence in Classroom Interactions," *Early Childhood Research Quarterly* 25, 1 (2010): 1-16; Amanda P. Williford, Michelle F. Maier, Jason T. Downer, Robert C. Pianta, and Carolee Howes, "Understanding How Children's Engagement and Teachers' Interactions Combine to Predict School Readiness," *Journal of Applied Developmental Psychology* 34, 6 (2013): 299-309; Amanda P. Williford, Jessica E. Vick Whittaker, Virginia E. Vitiello, and Jason T. Downer, "Children's Engagement Within the Preschool Classroom and Their Development of Self-Regulation," *Early Education and Development* 24, 2 (2013): 162-187.

[e]Stephanie M. Curenton, Iheoma U. Iruka, Marisha Humphries, Bryant Jensen, Tonia Durden, Shana E. Rochester, Jacqueline Sims, Jessica V. Whittaker, and Mable B. Kinzie, "Validity for the Assessing Classroom Sociocultural Equity Scale (ACSES) in Early Childhood Classrooms," *Early Education and Development* 31, 2 (2020): 284-303; Stephanie M. Curenton, Shana E. Rochester, Jacqueline Sims, Nneka Ibekwe-Okafor, Iheoma U. Iruka, A. G. García-Miranda, Jessica V. Whittaker, and Mable B. Kinzie, "Antiracism Defined as Equitable Sociocultural Interactions in Prekindergarten: Classroom Racial Composition Makes a Difference," unpublished manuscript (Boston: Boston University, 2021).

[f]Elena P. Soukakou, Pam J. Winton, Tracey A. West, John H. fSideris, and Lia M. Rucker, "Measuring the Quality of Inclusive Practices: Findings From the Inclusive Classroom Profile Pilot," *Journal of Early Intervention* 36, 3 (2015): 223-240.

[g]Meghan McCormick, JoAnn Hsueh, Mirjana Pralica, Catherine Snow, Christina Weiland, Amanda K. Weissman, Lillie Moffett, and Jason Sachs, "Is Skill Type the Key to the PreK Fadeout Puzzle? Differential Associations Between Enrollment in PreK and Constrained and Unconstrained Skills Across Kindergarten," *Child Development 92, 4* (2021): 599-620.

[h]Julie Sarama and Doug H. Clements, *Manual for Classroom Observation (COEMET)* (London: Routledge, 2009); Douglas H. Clements and Julie Sarama, "Early Childhood Mathematics Intervention," Science 333, 6045 (2011): 968-970.

[i]Miriam W. Smith, Joanne P. Brady, and Louisa Anastasopoulos, *User's Guide to the Early Language & Literacy Classroom Observation Pre-K Tool: ELLCO Pre-K* (Baltimore: Brookes Publishing Co., 2008); Carrie Ball and Maribeth Gettinger, "Monitoring Children's Growth in Early Literacy Skills: Effects of Feedback on Performance and Classroom Environments," *Education and Treatment of Children* 32, 2 (2009): 189-212; Ying Guo, Laura M. Justice, Joan N. Kaderavek, and Anita McGinty, "The Literacy Environment of Preschool Classrooms: Contributions to Children's Emergent Literacy Growth," *Journal of Research in Reading* 35, 3 (2012): 308-327; Barbara Jackson and Robert E. Larzelere, "The Impact of HeadsUp! Reading on Early Childhood Educators' Literacy Practices and Preschool Children's Literacy Skills," *Early Childhood Research Quarterly* 21, 2 (2006): 213-226; Miriam W. Smith and David K. Dickinson, *Early Language & Literacy Classroom Observation (ELLCO) Toolkit, Research Edition* (Baltimore: Brookes Publishing Co., 2002).

[j]Hanley Chiang, Elias Walsh, Timothy Shanahan, Claudia Gentile, Alyssa Maccarone, Tiffany Waits, Barbara Carlson, Samuel Rikoon, and Tracy Rimdzius, *An Exploration of Instructional Practices that Foster Language Development and Comprehension: Evidence from Prekindergarten through Grade 3 in Title I Schools* (Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2017).

Research on the ISI shows that learning experiences vary among children in the same classroom and across important subgroups.[30] In some studies, this measure has predicted children's gains in important skills, including in two studies that controlled for CLASS, though findings have been mixed.[31] Although the ISI is too resource-intensive to be used on a large scale or to provide data to individual teachers quickly enough to improve instruction, these findings highlight the potential importance of measuring children's individual classroom experiences.

Measuring Cultural Responsiveness and Inclusion: ACSES

Culturally responsive interactions and inclusion practices are understudied components of classroom quality. As shown in Table 2, the Assessing Classroom Sociocultural Equity Scale (ACSES), a novel measure, addresses this gap in the field by measuring whether classroom experiences are equitable for historically marginalized children of color. Specifically, trained, reliable observers rate how frequently teachers incorporate themes of social justice and non-stereotypical materials into their curricula; provide children of color with opportunities to develop critical thinking skills and build connections with them and encourage their participation in classroom discussions; address discipline in an equitable way; encourage children to share their experiences at home and in their communities; and offer personalized learning opportunities. Ratings can be scored with videos of approximately one hour of classroom time.

Although ACSES has been developed only recently, there is some initial evidence of its predictive validity from a sample of 105 children in 20 classrooms.[32] One of the five scales—Challenging Status Quo—was the most consistent predictor of children's gains, including predicted gains in math, execution function, and social competence. Equitable Discipline, another of the five scales, predicted gains in one skill (math), while Connections to Home Life predicted gains in two (executive function and competence). One scale—equitable learning opportunities—actually predicted slower growth in children's math and executive functioning. Relations between the scales and children's problem behavior were null. More research is needed to examine psychometric properties of this measure across systems and localities throughout the country.

Measuring Content: COEMET and the Boston Fidelity Measure

Table 2 also highlights six measures of content. One of these is the Classroom Observation of Early Mathematics Environment and Teaching (COEMET), which focuses specifically on the quality of math instruction in pre-K classrooms.[33] It captures the specific content of instruction (that is, numeral recognition and addition and subtraction strategies), the amount of time spent on math instruction, and the quality of math teaching strategies. A second measure of content is the Boston Fidelity Tool, which examines overall cognitive demand of instruction and content richness across all activities observed in the classroom. Both tools have been shown to predict gains in children's academic skills.[34] Interestingly, in the Boston Fidelity Tool study, the research team found that the tool predicted gains in children's math skill, even after controlling for CLASS scores. When examined on its own, the CLASS did not predict children's skills in any domain.[35] As illustrated in Figure 3, even when children started the pre-K year with the same level of math skills, those who received more content-rich instruction gained more during the academic year than their peers who received less content-rich instruction.

Figure 3. Gains in Pre-K Children's Math Skills in Classrooms with Varying Levels of Content-Rich Instruction

Applied Problems Raw Score

Time of school year

━━━ High content-rich instruction (1 SD above mean)

━━━ Low content-rich instruction (1 SD below mean)
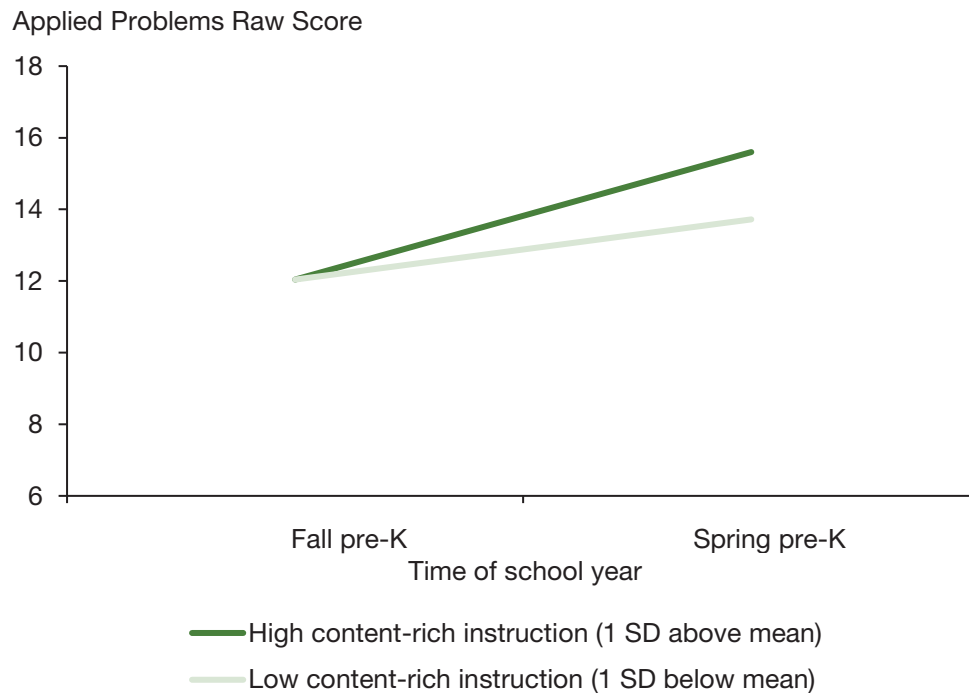
SOURCE: Michelle Maier, Meghan P. McCormick, Samantha Xia, JoAnn Hsueh, Christina Weiland, Abby Morales, Marina Boni, Melissa Tonachel, Jason Sachs, and Catherine Snow, "Content-Rich Instruction and Cognitive Demand in PreK: Using Systematic Observations to Predict Child Gains," *Early Childhood Research Quarterly* 60 (2022): 96-109.

Neither the COEMET nor the Boston Fidelity Tool has been widely used and neither would capture all important aspects of the quality of children's learning experiences when used in isolation. But both serve as interesting touch points for the field to identify promising new directions. More extensive use of measures that do capture instructional content across settings and localities could help guide curriculum development and implementation, an important direction given the evidence that high-quality content instruction can help promote more equitable learning opportunities and outcomes for children from marginalized groups.[36]

Future Directions for Strengthening Measures of Classroom Quality

As Table 2 illustrates, there is no single perfect measure of all of the important aspects of children's pre-K experiences. Although each existing measure has strengths, they are all fairly burdensome to administer and collect. They require considerable resources to train observers, code classroom instruction, analyze and

interpret data, and translate the results into actionable changes in professional development and practice. In this current context, here are five recommendations for the field of early learning and pre-K assessment moving forward. These recommendations are aimed at guiding the next generation of investments in measures of classroom quality.

1 **ADDRESS CRITICAL LIMITATIONS OF WIDELY USED MEASURES.** Prioritize strengthening existing tools and developing new ones that address the limitations of their predecessors, especially measures of children's individual learning experiences, cultural responsiveness, and content. Measuring these facets of classrooms is essential for identifying areas for improvement in the field and for addressing inequities in pre-K quality. The new versions of ECERS and CLASS that are currently being developed may hold particular promise, given the existing advantages of these measures described in this brief and their broad use in the field already

2 **LEVERAGE TECHNOLOGY AND NEW COLLECTION METHODS.** Increasingly, there are more ways to gather footage of classrooms using cameras and microphones.[37] New methods like machine learning can then be used to code videos and identify the most important dimensions of children's classroom experiences for their learning gains.[38] This approach might also include combining observation data with teacher and parent surveys and rich administrative data on operations, curriculum, staff qualifications, participation in professional development, bilingual supports, and special education supports. This work is in its infancy, especially in early education settings, but the approach eventually could be highly scalable and more cost-effective than current methods.

3 **FUND A COORDINATED MEASUREMENT DEVELOPMENT AND TESTING INITIATIVE.** Existing measures have many strengths, but their limitations call for new investments in measurement development and rigorous testing. The federal Preschool Curriculum Evaluation Research initiative is a model for this work. The 2002 initiative rigorously tested 14 different pre-K curricula and moved the field forward by demonstrating the efficacy of a content-specific curriculum rather than more general all-purpose curricula.[39] A similarly ambitious, careful era of development and testing is needed that is focused on how measures can be used at scale to improve access, equity, and opportunity.

4 **DESIGN MEASURES TO MONITOR AND PROMOTE CURRICULUM ALIGNMENT.** Programs tend to improve on dimensions that are measured.[40] Too often, instruction is not aligned as children move from pre-K into elementary school,[41] even though instructional alignment has been shown to help maintain early learning gains.[42] As new investments are made, attention should be paid to developing measures that can be used beyond pre-K to examine and improve children's instructional experiences throughout elementary school.

5 **MEASURES SHOULD BE OPEN ACCESS.** Investments in new measures should prioritize open access and a data-sharing infrastructure that makes the measures accessible across contexts and settings. Program resources are often scarce; existing measures can be outside the budget of many entities, especially in communities with low- and middle-incomes. Open access also facilitates more use of the same measures across contexts—a strength of current widely used measures that hopefully will characterize early childhood measurement using new tools as well.

With significant federal investment in public pre-K potentially on the horizon, accurately measuring pre-K quality on a large scale has perhaps never been more critical. Expanding access to quality early learning requires renewed attention to getting measurement of children's classroom experiences right. ■

Notes and References

1   Ajay Chaudry, Taryn Morrissey, Christina Weiland, and Hirokazu Yoshikawa, *Cradle to Kindergarten: A New Plan to Combat Inequality, 2nd edition* (New York: Russell Sage Foundation, 2021).

2   Christina Weiland, Lillie Moffet, Paola Guerrero Rosada, Amanda Weissman, K. Zhang, Michelle Maier, Catherine Snow, Meghan McCormick, JoAnn Hsueh, and Jason Sachs, "Is Child-Level Measurement the Key to Improving the Predictive Validity of Observational Measures in Early Education Classrooms?" unpublished paper (New York: MDRC, 2021).

3   Margaret Burchinal, "Measuring Early Care and Education Quality," *Child Development Perspectives* 12, 1 (2017): 3-9, https://doi.org/10.1111/cdep.12260; Christina Weiland, "Commentary: Pivoting to the 'How': Moving Preschool Policy, Practice, and Research Forward," *Early Childhood Research Quarterly* 45 (2018): 188–192, https://doi.org/10.1016/j.ecresq.2018.02.017.

4   Debby Cryer, Richard M. Clifford, and Thelma Harms, *Early Childhood Environment Rating Scale-Revised Edition* (New York: Teachers College Press, 2005); Thelma Harms, Richard Clifford, and Debby Cryer, *Early Childhood Environment Rating Scale Third Edition* (New York: Teachers College Press, 2014); Robert C. Pianta, Karen M. La Paro, and Bridget K. Hamre, *Classroom Assessment Scoring System™: Manual K-3* (Baltimore: Paul H. Brookes Publishing, 2008).

5   Daphna Bassok, Thomas S. Dee, and Scott Latham, "The Effects of Accountability Incentives in Early Childhood Education," *Journal of Policy Analysis and Management* 38, 4 (2019): 838-866, https://doi.org/10.1002/pam.22149; Daphna Bassok, Preston Magouirk, and Anna J. Markowitz, "Systemwide Quality Improvement in Early Childhood Education: Evidence from Louisiana," *AERA Open* 7 (2021), https://doi.org/10.1177/23328584211011610.

6   Bridget K. Hamre, Robert C. Pianta, Margaret Burchinal, Samuel Field, Jennifer LoCasale-Crouch, Jason T. Downer, Carollee Howes, Karen La Paro, and Catherine Scott-Little, "A Course on Effective Teacher-Child Interactions: Effects on Teacher Beliefs, Knowledge, and Observed Practice," *American Educational Research Journal* 49, 1 (2012): 88–123, https://doi.org/10.3102/0002831211434596; Robert C. Pianta, Bridget K. Hamre, Jason Downer, Margaret Burchinal, Amanda Williford, Jennifer LoCasale-Crouch, Carolle Howes, Karen La Paro, and Scott-Little, "Early Childhood Professional Development: Coaching and Coursework Effects on Indicators of Children's School Readiness," *Early Education and Development* 28, 8 (2017): 956–975, https://doi.org/10.1080/10409289.2017.1319783.

7   Ajay Chaudry, Taryn Morrissey, Christina Weiland, and Hirokazu Yoshikawa, *Cradle to Kindergarten: A New Plan to Combat Inequality* (New York: Russell Sage Foundation, 2017); Sara Anderson and Megan Mikesell, "Child Care Type, Access, and Quality in Rural Areas of the United States: A Review," *Early Child Development and Care* 189, 11 (2019): 1812-1826; Daphna Bassok and Eva Galdo, "Inequality in Preschool Quality? Community-Level Disparities in Access to High-Quality Learning Environments," *Early Education and Development* 27, 1 (2016): 128–144, https://doi.org/10.1080/10409289.2015.1057463; Scott Latham, Sean P. Corcoran, Carolyn Sattin-Bajaj, and Jennifer L. Jennings, "Racial Disparities in Pre-K Quality: Evidence From New York City's Universal Pre-K Program," *Educational Researcher* 50, 9 (2021): 607–617, https://doi.org/10.3102/0013189X211028214.

8   Mark Schneider, "Making Common Measures More Common," May (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, 2020).

9   The Build Initiative and Child Trends, "Quality Compendium: A Catalog and Comparison of Quality Improvement Systems," website: https://qualitycompendium.org, 2015.

10 Hirokazu Yoshikawa, Christina Weiland, Jeanne Brooks-Gunn, Margaret R. Burchinal, Linda M. Espinosa, William T. Gormley, Jens Ludwig, Katherine A. Magnuson, Deborah Phillips, and Martha J. Zaslow, *Investing in Our Future: The Evidence Base on Preschool Education* (Washington, DC: Foundation for Child Development, 2013).

11 Christina Weiland, Kchersti Ulvestad, Jason Sachs, and Hirokazu Yoshikawa, "Associations Between Classroom Quality and Children's Vocabulary and Executive Function Skills in an Urban Public Prekindergarten Program," *Early Childhood Research Quarterly* 28, 2 (2013): 199–209, https://doi.org/10.1016/j.ecresq.2012.12.002.

12 Ashley Brunsek, Michal Perlman, Olesya Falenchuk, Evelyn McMullen, Brooke Fletcher, and Prakesh S. Shah, "The Relationship Between the Early Childhood Environment Rating Scale and Its Revised Form and Child Outcomes: A Systematic Review and Meta-Analysis," *PLOS ONE* 12, 6 (2017), https://doi.org/10.1371/journal.pone.0178512; Margaret Burchinal, Lynne Vernon-Feagans, Virginia Vitiello, and Mark Greenberg, "Thresholds in the Association Between Child Care Quality and Child Outcomes in Rural Preschool Children," *Early Childhood Research Quarterly* 29, 1 (2014): 41-51, https://doi.org/10.1016/j.ecresq.2013.09.004; Jennifer K. Finders, Adassa Budrevich, Robert J. Duncan, David J. Purpura, James Elicker, and Sara A. Schmitt, "Variability in Preschool CLASS Scores and Children's School Readiness," *AERA Open* 7 (2021), https://doi.org/10.1177/23328584211038938; Michal Perlman, Olesya Falenchuk, Brooke Fletcher, Evelyn McMullen, Joseph Beyene, and Prakesh S. Shah, "A Systematic Review and Meta-Analysis of a Measure of Staff/Child Interaction Quality (the Classroom Assessment Scoring System) in Early Childhood Education and Care Settings and Child Outcomes," *PLOS ONE* 11, 12 (2016), https://doi.org/10.1371/journal.pone.0167660; Tyler W. Watts, Jill Gandhi, Deanna A. Ibrahim, Michael D. Masucci, and C. Cybele Raver, "The Chicago School Readiness Project: Examining the Long-Term Impacts of an Early Childhood Intervention," *PLOS ONE* 13, 7 (2018), https://doi.org/10.1371/journal.pone.0200144; Weiland, Ulvestad, Sachs, and Yoshikawa (2013).

13 Paola Guerrero-Rosada, Christina Weiland, Meghan McCormick, JoAnn Hsueh, Jason Sachs, Catherine Snow, and Michelle Maier, "Null Relations Between CLASS Scores and Gains in Children's Language, Math, and Executive Function Skills: A Replication and Extension Study," *Early Childhood Research Quarterly* 54 (2021): 1–12, https://doi.org/10.1016/j.ecresq.2020.07.009.

14 Nikki Aikens, Tutrang Nguyen, and Jessica Harding, *How Much Does the Pre-K CLASS Relate to Children's Readiness for School Skills? Early Childhood Literature Scan Brief* (Princeton, NJ: Mathematica, 2021). Aikens and her colleagues, in a summary of findings from recent studies, examined the strength of relations between CLASS scores and children's gains in subgroup analyses. They found that high levels of Instructional Support were stronger predictors of reading skills for children from families with higher incomes, native English speakers benefited more from higher Instructional and Emotional Support, and a composite CLASS measure predicted math gains for children from non-immigrant families, but not immigrant families.

15 Finders, Budrevich, Duncan, Purpura, Elicker, and Schmitt (2021); Kara M. Styck, Christopher J. Anthony, Lea E. Sandilos, and James C. DiPerna, "Examining Rater Effects on the Classroom Assessment Scoring System," *Child Development* 92, 3 (2021): 976–993, https://doi.org/10.1111/cdev.13460.

16 Anna-Katharina Praetorius, Christine Pauli, Kurt Reusser, Katrin Rakoczy, and Eckhard Klieme, "One Lesson is All You Need? Stability of Instructional Quality across Lessons," *Learning and Instruction* 31 (2014): 2–12, https://doi.org/10.1016/j.learninstruc.2013.12.002.

17 Margaret Burchinal, Kylie Garber, Tiffany Foster, Mary Bratsch-Hines, Ximena Franco, and Ellen Peisner-Feinberg, "Relating Early Care and Education Quality to Preschool Outcomes: The Same or Different Models for Different Outcomes?" *Early Childhood Research Quarterly* 55 (2021): 35-51, https://doi.org/10.1016/j.ecresq.2020.10.005; Carol M. Connor, Frederick J. Morrison, and Lisa Slominski, "Preschool Instruction and Children's Emergent Literacy Growth," *Journal of Educational Psychology* 98, 4 (2006): 665–689, https://doi.org/10.1037/0022-0663.98.4.665; Lillie Moffett and Frederick J. Morrison, "Off-Task Behavior in Kindergarten: Relations to Executive Function and Academic Achievement," *Journal of Educational Psychology* 112, 5 (2020): 938–955, https://doi.org/10.1037/edu0000397; Terri J. Sabol, Natalie L. Bohlmann, and Jason T. Downer, "Low-Income Ethnically Diverse Children's Engagement as a Predictor of School Readiness Above Preschool Classroom Quality," *Child Development* 89, 2 (2018): 556–576, https://doi.org/10.1111/cdev.12832; Virginia E. Vitiello, Leslie M. Booren, Jason T. Downer, and Amanda P. Williford, "Variation in Children's Classroom Engagement Throughout a Day in Preschool: Relations to Classroom and Child Factors," *Early Childhood Research Quarterly* 27, 2 (2012): 210–220, https://doi.org/10.1016/j.ecresq.2011.08.005.

18 Walter S. Gilliam, "Prekindergarteners Left Behind: Expulsion Rates in State Prekindergarten Systems," May (New York: Foundation for Child Development, 2005); Weiland et al. (2021).

19 Linda M. Espinosa and Michael L. López, "Assessment Considerations for Young English Language Learners Across Different Levels of Accountability," *National Early Childhood Accountability Task Force and First* 5 (2007); Jeanne L. Reid and Douglas D. Ready, "High-Quality Preschool: The Socioeconomic Composition of Preschool Classrooms and Children's Learning," *Early Education & Development* 24, 8 (2013): 1082–1111, https://doi.org/10.1080/10409289.2012.757519; Anna Rhoad-Drogalis and Laura M. Justice, "Is the Proportion of Children with Disabilities in Inclusive Preschool Programs Associated with Children's Achievement?" *Journal of Early Intervention* 42, 1 (2020): 83–96, https://doi.org/10.1177/1053815119873100; Christina Weiland, "Launching Preschool 2.0: A Road Map to High-Quality Public Programs at Scale," *Behavioral Science & Policy* 2, 1 (2016): 37–46, https://doi.org/10.1353/bsp.2016.0005.

20 National Association for the Education of Young Children "Advancing Equity in Early Childhood Education: A Position Statement of the National Association for the Education of Young Children," website: https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/advancingequitypositionstatement.pdf, 2019.

21 Susan H. Landry, Tricia A. Zucker, Jeffrey M. Williams, Emily C. Merz, Cathy L. Guttentag, and Heather B. Taylor, "Improving School Readiness of High-Risk Preschoolers: Combining High Quality Instructional Strategies with Responsive Training for Teachers and Parents," *Early Childhood Research Quarterly* 40 (2017): 38–51, https://doi.org/10.1016/j.ecresq.2016.12.001; Kent McIntosh, Calli B. Craft, and Leslie D. MacKay, "Perceived Cultural Responsiveness and Effectiveness of a Speech and Language Program for Indigenous Preschool Students," *Multicultural Learning and Teaching* 8, 1 (2013): 47–64, https://doi.org/10.1515/mlt-2012-0008; Geneva Gay, "Preparing for Culturally Responsive Teaching. *Journal of Teacher Education* 53, 2 (2002): 106–116.

22 Phillips et al. (2017); Yoshikawa et al. (2013); Deborah A. Phillips, Mark W. Lipsey, Kenneth A. Dodge, Ron Haskins, Daphna Bassok, Margaret R. Burchinal, Greg J. Duncan, Mark Dynarski, Katerine A. Magnuson, and Christina Weiland, *The Current State of Scientific Knowledge on Pre-Kindergarten Effects* (Washington, DC: Brookings Institution, 2017).

23 Douglas H. Clements and Julie Sarama, "Early Childhood Mathematics Intervention," *Science* 333, 6045 (2011): 968–970, https://doi.org/10.1126/science.1204537.

24 Pianta et al. (2017).

25 Early Childhood Learning and Knowledge Center, "National Overview of Grant Recipients CLASS® Scores by Year," website: https://eclkc.ohs.acf.hhs.gov/data-ongoing-monitoring/article/national-overview-grant-recipients-class-scores-year, 2021.

26 Ajay Chaudry, Taryn Morrissey, Christina Weiland, and Hirokazu Yoshikawa, *Cradle to Kindergarten: A New Plan to Combat Inequality*, *2nd edition* (New York: Russell Sage Foundation, 2021).

27 Karen L. Bierman, "50 Years of Research to Improve Preschool Curricula: Is there Progress?" February (Washington, DC: Monograph Matters, Society for Research in Child Development, 2021); Young Sun Joo, Katherine Magnuson, Greg J. Duncan, Holly S. Schindler, Hirokazu Yoshikawa, and Kathleen M. Ziol-Guest, "What Works in Early Childhood Education Programs? A Meta-Analysis of Preschool Enhancement Programs," *Early Education and Development* 31, 1 (2020): 1–26, https://doi.org/10.1080/10409289.2019.1624146; Yoshikawa et al. (2013).

28 Carol M. Connor, Frederick J. Morrison, Barry J. Fishman, Claire C. Ponitz, Stephanie Glasney, Phyllis S. Underwood, Shayne B. Piasta, Elizabeth C. Crowe, and Christopher Schatschneider, "The ISI Classroom Observation System: Examining the Literacy Instruction Provided to Individual Students," *Educational Researcher* 38, 2 (2009): 85-99, https://doi.org/10.3102/0013189X09332373.

29 Moffett and Morrison (2020); Weiland, Moffett, Guerrero Rosada, Weissman, Zhang, Maier, Snow, McCormick, Hsueh, and Sachs (2021).

30 Weiland et al. (2021).

31 Moffett and Morrison (2020); Weiland et al. (2021); Stephanie Al Otaiba, Carol Connor, Holly Lane, Marcia L Kosanovich, Chris Schatschneider, Allison K. Dyrlund, Melissa S. Miller, and Tyran L.Wright, "Reading First Kindergarten Classroom Instruction and Students' Growth in Phonological Awareness and Letter Naming-Decoding Fluency," *Journal of School Psychology* 46, 3 (2008): 281-314; Carol M. Connor, Frederick J. Morrison, and Leslie Katch, "Beyond the Reading Wards: Exploring the Effect of Child-Instruction Interactions on Growth in Early Reading," *Scientific Studies of Reading* 8, 4 (2004): 305-336, https://doi.org/10.1207/s1532799xssr0804_1; Connor, Morrison, and Slominski (2006); Moffett and Morrison (2020); Weiland et. al (2021).

32 Stephanie M. Curenton, Shana E. Rochester, Jacqueline Sims, Nneka Ibekwe-Okafor, Iheoma U. Iruka, A. G. García-Miranda, Jessica V. Whittaker, and Mable B. Kinzie, "Antiracism Defined as Equitable Sociocultural Interactions in Prekindergarten: Classroom Racial Composition Makes a Difference," unpublished manuscript (Boston: Boston University, 2021).

33 Julie Sarama and Doug H. Clements *Manual for Classroom Observation (COEMET),* (London: Routledge, 2009).

34 Meghan P. McCormick, Shira K. Mattera, Michelle F. Maier, Samantha Xia, Robin Jacob, and Pamela A. Morris, "Different S, Different Patterns of Impacts: Effects of a Pre-K Math Intervention in a Mixed-Delivery System," *Early Childhood Research Quarterly* 58 (2022): 136–154, https://doi.org/10.1016/j.ecresq.2021.08.005; Douglas H. Clements, Julie Sarama, Christopher B. Wolfe, and Marry Elaine Spitler, "Longitudinal Evaluation of a Scale-Up Model for Teaching Mathematics with Trajectories and Technologies: Persistence of Effects in the Third Year," *American Educational Research Journal* 50, 4 (2013): 812-850, https://doi.org/10.3102/0002831212469270; McCormick et al. (2022).

35 Guerrero-Rosada et al. (2021); McCormick et al. (2022).

36 Douglas Frye, Arthur J. Baroody, Margaret Burchinal, Sharon M. Carver, Nancy C. Jordan, and Judy McDowell, *Teaching Math to Young Children: A Practice Guide* (Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2013); Barbara Foorman, Nicholas Beyler, Kelley Borradaile, Michael Coyne, Carolyn A. Denton, Joseph Dimino, Joshua Furgeson, Lynda Hayes, Juliette Henke, and Laura Justice, *Foundational Skills to Support Reading for Understanding in Kindergarten through 3rd Grade. Educator's Practice Guide* (Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2016).

37 Maureen J. Myrtil, M. Tzu-Jung Lin, Jing Chen, Kelly M. Purtell, Laura Justice, Jessica Logan, and Allie Hamilton, "Pros and (Con)flict: Using Head-Mounted Cameras to Identify Teachers' Roles in Intervening in Conflict Among Preschool Children," *Early Childhood Research Quarterly* 55 (2021): 230–241, https://doi.org/10.1016/j.ecresq.2020.11.011; Yuanyuan Wang, Rondeline Williams, Laura Dilley, and Derek M. Houston, "A Meta-Analysis of the Predictability of LENA^{TM} Automated Measures for Child Language Development," *Developmental Review* 57 (2020),https://dx.doi.org/10.1016%2Fj.dr.2020.100921.

38 Patricia Goldberg, Ömer Sümer, Kathleen Stürmer, Wolfgang Wagner, Richard Göllner, Peter Gerjets, Enkelejda Kasneci, and Ulrich Trautwein, "Attentive or Not? Toward a Machine Learning Approach to Assessing Students' Visible Engagement in Classroom Instruction," *Educational Psychology Review* 33, 1 (2021): 27–49; Sean Kelly, Andrew M. Olney, Patrick Donnelly, Martin Nystrand, and Sidney K. D'Mello, "Automatically Measuring Question Authenticity in Real-World Classrooms," *Educational Researcher* 47, 7 (2018): 451–464; Anusha James, Mohan Kashyap, Yi Han Victoria Chua, Tomasz Maszczyk, Ana Moreno Núñez, Rebecca Bull, and Justin Dauwels, "Inferring the Climate in Classrooms from Audio and Video Recordings: A Machine Learning Approach," paper presented at the IEEE International Conference on Teaching, Assessment, and Learning for Engineering, Wollongong, Australia (December 4-7, 2018).

39 Jade Marcus Jenkins, and Greg J. Duncan, "Do Pre-Kindergarten Curricula Matter?" pages 37-43 in *The Current State of Scientific Knowledge on Pre-Kindergarten Effects* (Washington, DC: Brookings Institution, 2017).

40 Bassok, Dee, and Latham (2019).

41 Meghan P. McCormick, Christina Weiland, JoAnn Hsueh, Michelle Maier, Rama Hagos, Catherine Snow, Nicole Leacock, and Laura Schick, "Promoting Content-Enriched Alignment across the Early Grades: A Study of Policies & Practices in the Boston Public Schools," *Early Childhood Research Quarterly* 52 (2020): 57–73, https://doi.org/10.1016/j.ecresq.2019.06.012.

42 Clements, Sarama, Wolfe, and Spitler (2013); Robin Jacob, Anna Erickson, and Shira Mattera, "Evaluating the Impact of Small Group Supplemental Math Enrichment in Kindergarten," *Journal of Research on Educational Effectiveness* 13, 3 (2020): 381–407.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY