CHANGES IN PRESCHOOL CLASSROOM QUALITY

The Timing of Rating Matters

FEBRUARY 2023

AMY TAUB
MICHELLE F. MAIER
JOANN HSUEH

Many early care and education (ECE) institutions—including Head Start, state quality rating and improvement systems (QRIS), and school district pre-K programs—use classroom quality observations in their quality improvement or monitoring processes. As many states expand their ECE programs, they grapple with questions about how best to use data from quality observations, which are typically conducted once a year at most. This brief presents results from a study that examined levels and patterns of change in classroom quality over a school year, and the way those changes vary by different quality dimensions. The study is based on data from the Expanding Children's Early Learning (ExCEL) Quality project, which assessed two different approaches to improving classroom quality. The findings suggest that practitioners and policymakers should carefully consider the timing and number of quality ratings when making decisions about how to improve programs and monitor processes.

This brief examines how biweekly ratings by teachers' coaches of three dimensions of classroom quality—structural, interactional, and instructional quality—changed over the course of the study (which occurred from September 2019 through March 2020). It addresses two research questions:

- How do different dimensions of quality change over time?
- How do changes in the three dimensions of quality differ based on the curricular and professional development approach?

The larger ExCEL Quality study was designed to examine how curriculum combined with professional development helps teachers improve classroom quality and maximize improvements in children's outcomes. It is being conducted by MDRC and its partners, MEF Associates and RTI International, and is sponsored by Arnold Ventures.

CLASSROOM QUALITY IN PRESCHOOL

Despite widespread agreement that classroom quality affects children's outcomes,[1] the ECE field has not yet established a single definition for "high quality." Classroom quality has typically been examined in terms of two dimensions: structural quality and process quality. The ExCEL Quality study builds on this concept of quality by examining three separate but related dimensions of quality:[2]

- **STRUCTURAL QUALITY** refers to the physical aspects of a classroom, such as the materials and toys it contains, its cleanliness and safety, and its arrangement of furniture.

- **INTERACTIONAL QUALITY** refers to the quality of children's interactions with teachers and other children in the classroom. It includes whether the classroom climate is positive, responsive, and predictable; whether teachers are warm and sensitive; and how well the classroom is managed and organized.

- **INSTRUCTIONAL QUALITY** refers to which skills domains are taught in the classroom and how they are taught—that is, whether high-quality instructional practices that promote children's cognitive development are used.

PATTERNS OF CHANGE IN PRESCHOOL CLASSROOM QUALITY

Only a few studies have examined whether and how ECE classroom quality changes over the course of a school year. Some studies show increases in quality from fall to spring, but others suggest quality may fluctuate within or across school years.[3] One study of pre-K classrooms measured quality with the widely used Classroom Assessment Scoring System (CLASS) and found that measures of interactional and instructional quality generally increased over the pre-K year.[4] But the researchers also found that ratings for the dimensions of quality peaked at different times.[5] For example, instructional quality (as measured by the CLASS Instructional Support domain) was highest in December and January, whereas one aspect of interactional quality (as measured by the CLASS Classroom Organization domain) was highest in February through April. These findings show that it is important to understand how the dimensions of classroom quality might vary at different points during a school year, especially when quality ratings are used for accountability purposes.

Existing studies on this topic are limited; most have small samples, collect classroom quality data only a couple of times a year, and measure classroom quality solely with the CLASS.[6] The ExCEL Quality study aims to extend this research base by examining how three distinct dimensions of classroom quality may change over a school year, using more frequent, biweekly quality ratings from teachers' coaches.

CLASSROOM QUALITY AND THE ECE CURRICULUM

Providing a developmentally appropriate early childhood curriculum—and combining it with professional development for teachers—is a promising way to improve classroom quality and children's early learning experiences.[7] The goals, scope, materials, and activity settings (for example, small or large group settings) of ECE curricula vary, which means that classrooms using different curricula likely also provide different classroom activities.[8] In turn, the curriculum being used in a classroom is likely to influence classroom quality and how it may change over a school year. This brief examines whether changes in classroom quality over time are related to the curriculum implemented within the classroom.

STUDY DESIGN

The ExCEL Quality study was a cluster randomized controlled study that collected biweekly ratings of classroom quality from teachers' coaches.[9] The goal was to assess the effectiveness of two different curricular and professional development approaches to improving various dimensions of classroom quality:

- A **WHOLE-CHILD, GLOBAL APPROACH** that targets structural and interactional quality, which used the sixth edition of the Creative Curriculum®. This curriculum targets structural quality by focusing on the classroom physical set-up (for example, the arrangement of different learning areas) and materials, and it targets interactional quality by focusing on classroom routines, teachers' emotional sensitivity, positive behavior management, and child-led activities.

- An **INTEGRATED, DOMAIN-SPECIFIC APPROACH** that targets instructional quality, which used the first edition of Connect4Learning®. This curriculum uses a specific scope and sequence—that is, the curricular activities target a specific set of learning domains and are placed in a particular order in which they should be taught. Connect4Learning emphasizes instructionally rich interactions between teachers and children, such as asking open-ended questions, facilitating in-depth discussions with children, and asking children to use higher-order thinking skills such as comparing and contrasting different things.

ExCEL Quality took place in 53 ECE centers (located in four U.S. cities) that serve children from families with low incomes. The centers were randomly assigned to implement one of the two intervention approaches described above, or they were assigned to a preschool-as-usual control group. Eighteen centers were assigned to each of the two intervention approaches (for 36 total intervention centers); 17 centers were assigned to the control group.

Lead and assistant teachers implementing the two intervention approaches received the following professional development supports to help implement the assigned curriculum:

- Summer and academic year training were provided by the curriculum developers.[10]

- Biweekly in-person coaching sessions, consisting of a two-hour classroom observation and a one-hour meeting between the coach and teacher(s), were provided by coaches hired by the ExCEL Quality study's local coaching partners.

Because the current study examines classroom quality data that coaches reported on an ongoing basis, this brief focuses only on classrooms assigned to one of the two intervention approaches.

Classrooms were eligible for the study if they served 4-year-old children or mixed-age groups of 3- and 4-year-old children. In total, the centers implementing the two intervention approaches comprised 95 eligible classrooms that served 1,527 children.[11] The study sample included 21 communi-

ty-based child care centers and 15 Head Start centers.[12] Table 1 provides more information about the study sample and its breakdown across intervention approach and setting.

TABLE 1

THE STUDY SAMPLE

SETTING	INTEGRATED, DOMAIN-SPECIFIC APPROACH USED	WHOLE-CHILD, GLOBAL APPROACH USED
Head Start setting	7 centers 22 classrooms	8 centers 27 classrooms
Community-based child care setting	11 centers 23 classrooms	10 centers 23 classrooms

Fifteen of the centers had received professional development supports the previous year through a pilot study for the Variations in Implementation of Quality Interventions project (VIQI).[13] Because ExCEL Quality uses the same two intervention approaches as VIQI, these centers received the same curricular and professional development model over two years. The other 21 centers were newly recruited and only received one year of professional development supports.

Random assignment took place in July 2019, and professional development supports and data collection were scheduled to continue through the spring of 2020. Due to the COVID-19 pandemic, most centers suspended in-person operations in March 2020. This brief focuses on the activities that were conducted prior to the closures.

Coach Ratings of Classroom Quality

This study used a novel quality rating system to capture classroom quality on a biweekly basis in a cost-effective manner. Coaches assessed quality using a web survey called a "coach log." They reported whether they observed different aspects of classroom quality in a series of yes-or-no questions. Despite the simplicity of the questions, the rating system was sensitive enough to capture variation in quality. That is, the questions captured aspects of quality that were present in some of the classrooms some of the time. This variation makes it possible to use these questions to look at differences between classrooms at each time point and within classrooms over time.

Throughout the study, teachers in each classroom implementing an intervention approach had a coaching session every other week. During these sessions, the coach observed the classroom and met with the teacher(s). Afterward, the coach completed the coach log—which took approximately 15 minutes—to report what occurred during the session. The purpose of the coach log was twofold: It served as a rich data source to investigate how classrooms functioned over the course of the year, and it helped the study team see what technical assistance and professional development supports were needed during the study period.

One section of the log asked coaches to report whether certain indicators of classroom quality were present during the classroom observation portion of the coaching session. (See Box 1 for examples of these classroom quality indicators.) The percentage of the indicators that were observed during the coaching sessions were used to measure three dimensions of quality: structural, interactional, and instructional quality. [14] This brief examines the classroom quality indicators over the course of the study.

BOX 1

MEASURES

Coaches reported when they saw indicators of classroom quality during their biweekly coaching sessions. The three dimensions of quality were measured by looking at the percentage of indicators that were present.

STRUCTURAL QUALITY was measured using four indicators, including "Areas were arranged so children can easily get around and access materials" and "Classroom was neat and clean."

INTERACTIONAL QUALITY was measured using 19 indicators, including "Transitions between activities were used as opportunities to teach skills and concepts," "Children knew the routines and expectations in the classroom very well," and "Children talked to teachers/adults in the classroom about a variety of topics."

INSTRUCTIONAL QUALITY was measured using 15 indicators, including "Asking children 'why' and 'how' questions," "Changing an activity or portion of an activity to make it easier for individual children who need more support," and "Using and defining advanced vocabulary words (such as words with more than one meaning or words that are not used frequently)."

Coach log data are available in two-week reporting periods from September 2019 to early March 2020. In total, coaches completed 1,043 coach logs for 95 classrooms. For each classroom, 7 to 13 logs were completed, with an average of 11 logs. Classrooms using the two intervention approaches had approximately the same average number of logs per classroom.

Coaches were trained on how to complete the coach logs during their onboarding process. Coaches were taught study-specific definitions of terms used in the logs to ensure that they detected and reported indicators consistently. They also practiced using the logs during training and had to pass a certification test on the terminology used in the logs. In the middle of the year, coaches were retrained on how to use the logs and reviewed the terminology again.

The Study Sample

On average, lead teachers were 44 years of age; their ages ranged from 21 to 68 years. They had been teaching preschool for an average of 12 years, with a range of 1 year to 40 years. Fifty-nine percent of teachers had a bachelor's degree or higher. Forty-eight percent of the teachers were non-Hispanic

Black, 32 percent were non-Hispanic White, and 15 percent were Hispanic.[15] On average, 16 children were enrolled in the study classrooms; enrollment ranged from 6 to 26 children.

Children in the classrooms were about 3.5 years old (ages ranged from 2 to 6 years old) at the beginning of the school year. Fifty-two percent of the children were female. Forty-three percent of the children were non-Hispanic Black, 38 percent were Hispanic, 11 percent were non-Hispanic White, and 8 percent were multiracial or another race. Parents were 32 years of age on average (their ages ranged from 18 to 83 years). Parents' most frequent level of educational attainment was a high school diploma or GED (63 percent); 23 percent of parents had at least some college experience.[16]

EXAMINING CHANGE OVER TIME

The research team used growth models to analyze changes in classroom quality over time. These models describe the average shape of the change in each of the quality dimensions from September 2019 to March 2020. Unlike looking at the simple averages for each time period, these models find a curve, or a smooth line with up to two bends, that best approximates the average change over time. The team examined multiple types of models for each quality dimension, allowing the curve to have zero, one, or two bends. This brief presents results from the best fitting model. (See Box 2 for more details on the analyses, including how the "best fitting model" was determined.) After finding the best-fitting growth model for the overall sample, the team assessed whether those trajectories differed by intervention approach.

BOX 2
ANALYTIC APPROACH

Growth models for each of the three dimensions of quality were conducted using hierarchical linear modeling (HLM) in R (lme4 package).

Because coaching sessions were supposed to occur once every two weeks, time was defined as the number of two-week coaching periods that had passed since the beginning of the study. Time was centered at the beginning of the school year (the first period was coded as 0, the second as 1, and so on). This means that the model intercepts (depicted in the figures as the *y*-axes) can be interpreted as the average quality indicator ratings at the beginning of the study.

Models were run iteratively, beginning with an empty model (no predictors) and adding polynomial terms for time [linear, quadratic ($time^2$), and cubic ($time^3$)], one at a time. Final models were selected using forward elimination—the results shown in this brief reflect the highest-order polynomial that was statistically significant for that measure. For example, if $time^2$ was statistically significant when it was the highest polynomial, but $time^3$ was not, the results from the model without $time^3$ are shown here. Covariates and interaction terms were only added to the models after the shape of the trajectory had been established.

(continued)

BOX 2 (continued)

Differences between the two intervention approaches were analyzed using interaction terms. Models included an interaction between the random assignment indicator and each of the terms for time. These interactions allowed the team to assess whether the trajectories were different for the subgroups based on curricular model. Because time was centered at the beginning of the school year, the team could also assess whether the differences between the intervention approach subgroups were larger than would be expected by chance at the beginning of the study.

Models control for a variety of classroom characteristics, such as class size, child and family demographic characteristics, and average child vocabulary scores on the auditory comprehension subscale of the Preschool Language Scales. Adjusting for covariates increases the precision of the estimates and produces findings that can be interpreted over and above differences in classroom characteristics (that is, differences between groups cannot be attributed to differences in classroom composition). The results were robust to sensitivity checks, including different sets of covariates, methods of imputing covariates, and elimination of coaching periods with lower response rates.

FINDINGS

- **Each dimension of quality had a different pattern of change over the study period.**

Figure 1 shows the average rating for the three dimensions of quality at each point in time, as well as the patterns of change predicted by the growth models. The raw average is shown as a dot, and the patterns of change are shown as a line. The vertical axis shows the percentage of quality indicators observed by coaches, and the horizontal axis represents each reporting period, from September to March. The shading around the lines represents a 90 percent confidence interval.[17]

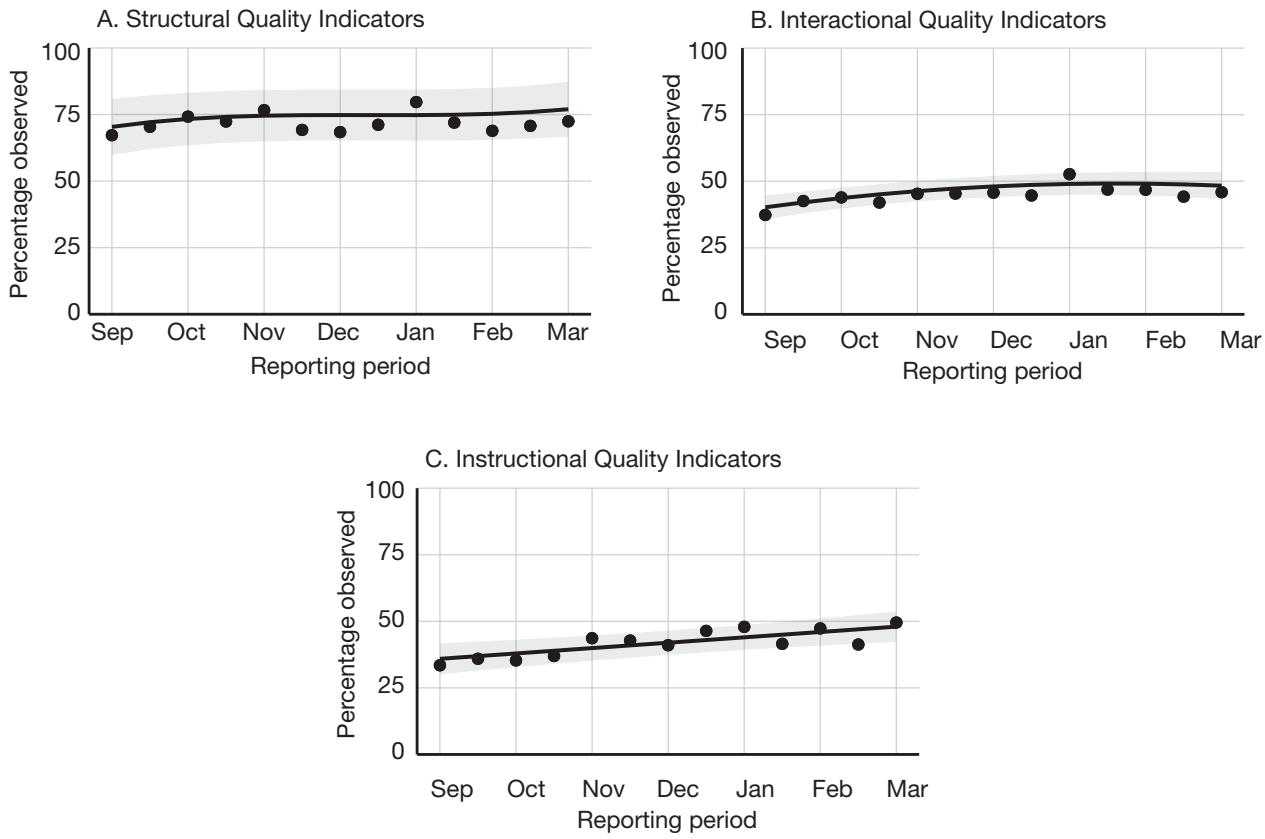
STRUCTURAL QUALITY. As shown in Figure 1, Model A, classrooms began the year with high structural quality (three-fourths of the structural quality indicators observed, on average) and did not exhibit statistically significant growth or change over time.

INTERACTIONAL QUALITY. As shown in Figure 1, Model B, coaches observed an average of about two-fifths of the interactional quality indicators in classrooms at the beginning of the year. The interactional quality ratings increased through the winter and peaked with coaches observing about half of the indicators in early February. The number of observed indicators declined slightly in early March.

INSTRUCTIONAL QUALITY. As shown in Figure 1, Model C, coaches observed about one-third of instructional quality indicators at the beginning of the school year, on average. The instructional quality ratings increased steadily throughout the study period, with coaches observing, on average, about half of the indicators in March.

FIGURE 1

AVERAGE PERCENTAGE OF QUALITY INDICATORS OBSERVED OVER TIME



SOURCE: ExCEL Quality Coach Log.

NOTES: The dots represent the raw means and the lines represent the means predicted by the growth model.

Structural quality indicators include features of the classroom set up, such as the arrangement of different areas and availability of materials. Interactional quality indicators include varied conversations between children and teachers, classroom management, and routines. Instructional quality indicators include scaffolding, differentiated instruction, elaboration, and other types of interactions that facilitate children's learning.

Model A includes linear, quadratic, and cubic terms for time; Model B includes linear and quadratic terms for time; and Model C includes a linear term for time.

The results of all three models are adjusted for classroom composition, class size, and average child vocabulary scores on the auditory comprehension subscale of the Preschool Language Scales.

The shaded area depicts a 90 percent confidence interval.

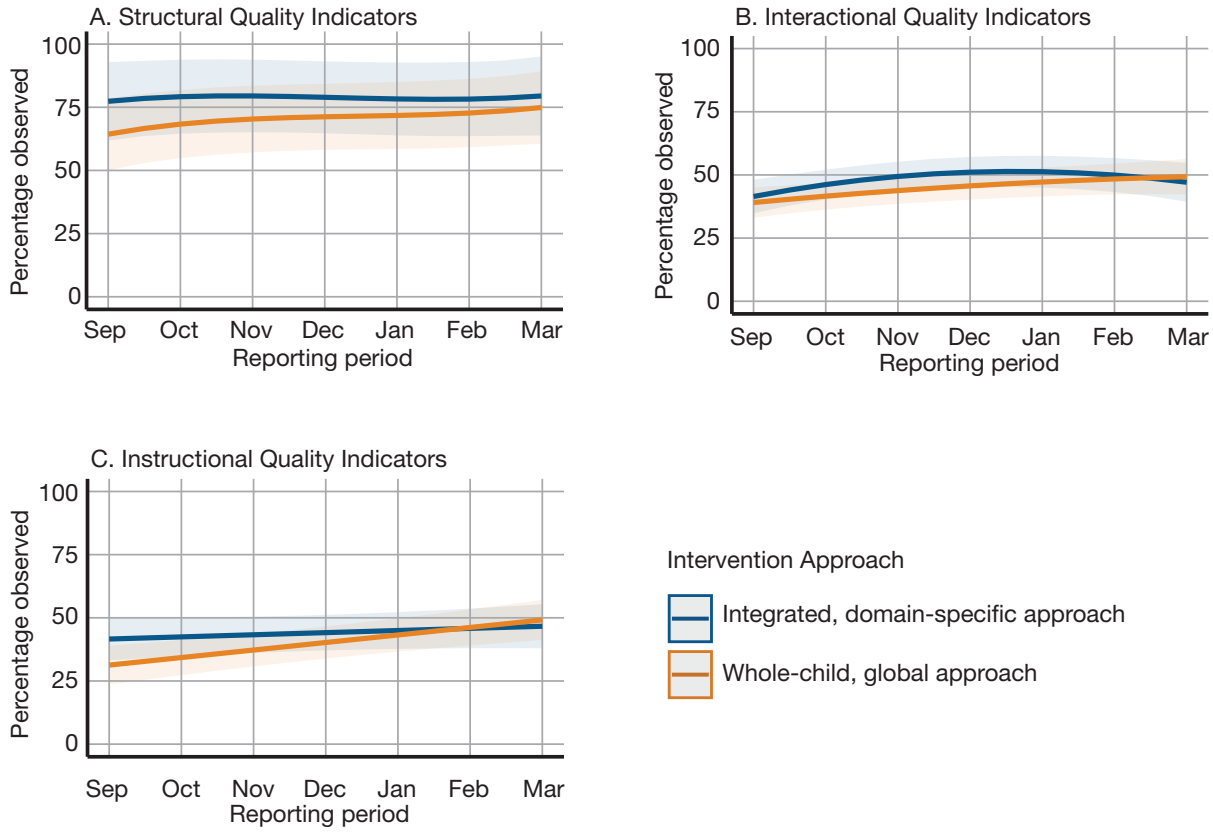
These findings are in line with prior research showing that different dimensions of quality may peak at different times during a school year.[18] Interactional quality appeared to peak in February, but instructional quality increased over time. In contrast, structural quality ratings remained consistent. Overall, changes in quality over the course of the study were modest in magnitude.

- **Patterns of change varied by intervention approach.**

Figure 2 shows the way the three dimensions of quality for each intervention approach change over time. The vertical axis shows the percentage of quality indicators coaches observed, and the horizontal axis shows each reporting period, from September to March. The curve for the integrated, domain-specific approach is shown in blue and the curve for the whole-child, global approach is shown in orange. The shading around the lines represents a 90 percent confidence interval.

FIGURE 2

PREDICTED AVERAGE PERCENTAGE OF QUALITY INDICATORS OBSERVED OVER TIME, BY INTERVENTION APPROACH



SOURCE: ExCEL Quality Coach Log.

NOTES: Structural quality indicators include features of the classroom set up, such as the arrangement of different areas and availability of materials. Interactional quality indicators include varied conversations between children and teachers, classroom management, and routines. Instructional quality indicators include scaffolding, differentiated instruction, elaboration, and other types of interactions that facilitate children's learning.

Model A includes linear, quadratic, and cubic terms for time; Model B includes linear and quadratic terms for time; and Model C includes a linear term for time.

The results of all three models are adjusted for classroom composition, class size, and average child vocabulary scores on the auditory comprehension subscale of the Preschool Language Scales.

The shaded area depicts a 90 percent confidence interval.

STRUCTURAL QUALITY. As shown in Figure 2, Model A, classrooms using both intervention approaches began the year with high structural quality ratings and did not exhibit any growth or change over time. Although more quality indicators were observed in the classrooms using the integrated, domain-specific approach throughout the study period, this difference was not statistically significant.

INTERACTIONAL QUALITY. As shown in Figure 2, Model B, in classrooms using the integrated, domain-specific approach, interactional quality ratings rose in the fall and early winter and declined in late February and early March. In contrast, interactional quality ratings in classrooms using the whole-child, global approach grew steadily through March. Although the two groups had different trajectories, their ratings were similar at the beginning and end of the study.[19]

INSTRUCTIONAL QUALITY. As shown in Figure 2, Model C, classrooms using the integrated, domain-specific approach had higher instructional quality ratings at the beginning of the year than classrooms using the whole-child, global approach.[20] However, classrooms using the integrated, domain-specific approach exhibited a slower rate of growth, so classrooms using both intervention approaches had a similar level of instructional quality by early March.[21]

These findings show that interactional and instructional quality ratings have different trajectories in classrooms implementing different intervention approaches.

CONCLUSION

The results show that a classroom's quality rating may depend on the time of year the classroom is observed. Moreover, classrooms that implement different intervention approaches or curricula may experience different fluctuations in classroom quality over the course of the school year. These findings have implications for how practitioners and policymakers conduct classroom observations focused on assessing quality, especially when quality information is used to inform professional development, program improvement, or program funding levels.

First, consistently conducting quality observations at the same general time across classrooms can help ensure that ratings are directly comparable to one another—that the ratings capture actual classroom quality rather than differences that may occur due to the time of year.

Second, collecting data at multiple time points throughout the year may create a clearer and more stable estimate of actual classroom quality compared with collecting data at a single time point. From logistical and resource perspectives, it is challenging to conduct multiple quality observations across classrooms at similar points in time, and to conduct multiple quality observations for each classroom during each school year, especially when hiring neutral third parties as observers. However, this study and others underscore that it is critical to consider the timing of observations and the number collected, particularly when the observations have high stakes for programs.[22] The coach ratings

used in this study offer a cost-effective, alternative method for collecting quality ratings on a more frequent basis.

Third, the way classroom quality trajectories differ can help inform the content and sequence of topics to discuss with teachers as part of their training and coaching. In the fall (early in the school year), professional development supports could focus on improving structural features of the classroom—such as how the room is arranged and which materials and manipulatives (objects like toys and counting cubes used by children for learning purposes) are available—because those aspects of quality are likely to stay in place for the rest of the year. In the spring (later in the school year), professional development supports could revisit aspects of interactional quality—such as the responsivity and sensitivity of teachers' interactions with children—since this study suggests interactional quality may start to decline around March. Finally, professional development supports could focus on instructional quality throughout the year, since it appears to continuously improve over time.

There are several considerations that should be noted. First, due to the COVID-19 pandemic, the research team was only able to collect data through early March. Classroom quality trajectories may have been different if the research team was able to extend data collection through the end of a typical school year. Second, the ratings used in these analyses were made by coaches who had a working relationship with the teachers they were monitoring. Neutral observers may have rated the classrooms differently. Third, since the quality ratings used in this study were from a novel tool, it is not clear whether the observed changes over time are large enough to have meaningful effects on children's experiences. Finally, these patterns were found in classrooms that were receiving professional development supports while implementing a new curriculum. Classroom quality trajectories may differ depending on whether teachers are implementing a familiar curriculum and receiving a high level of support.

What's Next?

Forthcoming briefs about the ExCEL Quality study will discuss additional implementation findings and information about how programs and families fared early in the COVID-19 pandemic. The research team also followed the study sample through the 2020–2021 school year to understand how centers navigated the pandemic and the extent to which they continued to implement the intervention approaches without study-provided supports. Findings from the 2020–2021 school year will also be highlighted in future publications.

NOTES AND REFERENCES

1 Margaret Burchinal, Carollee Howes, Robert Pianta, Donna Bryant, Diane Early, Richard Clifford, and Oscar Barbarin, "Predicting Child Outcomes at the End of Kindergarten from the Quality of Pre-Kindergarten Teacher–Child Interactions and Instruction," *Applied Developmental Science* 12, 3 (2008): 140–153; Margaret Burchinal, Martha Zaslow, and Louisa Tarullo, "Quality Thresholds, Features, and Dosage in Early Care and Education: Secondary Data Analyses of Child Outcomes," *The Society for Research in Child Development* 81, 2 (2016): 1–128.

2 Michelle F. Maier, JoAnn Hsueh, and Meghan McCormick, *Rethinking Classroom Quality: What We Know and What We Are Learning* (New York: MDRC, 2020).

3 Martha Buell, Myae Han, and Carol Vukelich, "Factors Affecting Variance in Classroom Assessment Scoring System Scores: Season, Context, and Classroom Composition," *Early Child Development and Care* 187, 11 (2017): 1–14; Anne H. Cash and Robert C. Pianta, "The Role of Scheduling in Observing Teacher-Child Interactions," *School Psychology Review* 43, 4 (2014): 428–449; Jill Gandhi, C. Cybele Raver, Rachel M. Abenavoli, Pamela A. Morris, and Lauren M. Meyer, "Variations in Pre-Kindergarten Classroom Quality Ratings Across the School Year: Observation Ratings from New York City's Pre-K for All," *Early Education and Development* 32, 8 (2020): 1,174–1,193; Susanne Kuger, Katharina Kluczniok, David Kaplan, and Hans-Guenther Rossbach, "Stability and Patterns of Classroom Quality in German Early Childhood Education and Care," *School Effectiveness and School Improvement* 27, 3 (2016): 418–440; Lars-Erik Malmberg, Hazel Hagger, Katharine Burn, Trevor Mutton, and Helen Colls, "Observed Classroom Quality During Teacher Education and Two Years of Professional Practice," *Journal of Educational Psychology* 102, 4 (2010): 916–932.

4 Observers using the CLASS rate classrooms in 10 quality dimensions, which are averaged to calculate scores for three domains: classroom organization (which captures aspects of interactional quality related to behavior management and learning formats), emotional support (which captures aspects of interactional quality related to classroom climate and teacher sensitivity), and instructional support (which captures aspects of instructional quality).

5 Cash and Pianta (2014).

6 Because previous studies have used the same observational tool, it is not clear whether their findings can be generalized to other classroom observation tools or to aspects of classroom quality that are not captured in the CLASS.

7 For example, Donna Bryant, Patricia W. Wesley, Margaret Burchinal, John Sideris, Karen Taylor, Carla Fenson, and Iheoma U. Iruka, *The QUINCE-PFI Study: An Evaluation of a Promising Model for Child Care Provider Training: Final Report* (Chapel Hill, NC: FPG Child Development Institute, 2009); Douglas H. Clements, Julie Sarama, Mary Elaine Spitler, Alissa A. Lange, and Christopher B. Wolfe, "Mathematics Learned by Young Children in an Intervention Based on Learning Trajectories: A Large-Scale Cluster Randomized Trial," *Journal for Research in Mathematics Education* 42, 2 (2011): 127–166; Kerry G. Hofer, Mark W. Lipsey, Nianbo Dong, and Dale C. Farran, "Results of the Early Math Project: Scale-Up Cross-Site Results," working paper (Nashville, TN: Vanderbilt University, Peabody Research Institute, 2013); Lisa G. Klein and Deanna S. Gomby, "A Synthesis of Federally Funded Studies on School Readiness: What Are We Learning About Professional Development?" working paper (Washington, DC: U.S. Department of Health and Human Services, 2008); Pamela A. Morris, Shira K. Mattera, and Michelle F. Maier, *Making Pre-K Count: Improving Math Instruction in New York City* (New York: MDRC, 2016); Pamela Morris, Shira K. Mattera, Nina Castells, Michael Bangser, Karen Bierman, and Cybele Raver, *Impact Findings from the Head Start CARES Demonstration: National Evaluation of Three Approaches to Improving Preschoolers' Social and Emotional Competence*, OPRE Report 2014-44 (Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2014); Hirokazu Yoshikawa, Diana Leyva, Catherine E. Snow, Ernesto Treviño, M. Clara Barata, Christina Weiland, Celia J. Gomez, Lorenzo Moreno, Andrea Rolla, Nikhit D'Sa, and Mary Catherine Arbour, "Experimental Impacts of a Teacher Professional Development Program in Chile on Preschool Classroom Quality and Child Outcomes," *Developmental Psychology* 51, 3 (2015): 309–322.

8 Jade M. Jenkins, Anamarie Auger Whitaker, Tutrang Nguyen, and Winnie Yu, "Distinctions Without a Difference? Preschool Curricula and Children's Development," *Journal of Research on Educational Effectiveness* 12, 3 (2016): 514–549.

9 In a cluster randomized controlled study, groups—rather than individuals—are randomly assigned to intervention groups. The ExCEL Quality team randomly assigned ECE centers within blocks—participating ECE centers were grouped by locality and whether they were Head Start centers or community-based child care centers before they were randomly assigned to one of the three groups.

10 The amount of training was prescribed by the curriculum developers and varied for the two approaches. Classrooms using the whole-child, global approach received two days of summer training and three days of academic year training. Classrooms using the integrated, domain-specific approach received three days of summer training and four and a half days of academic year training.

11 Three additional classrooms were identified by participating centers before random assignment, but they did not open and were not part of the study sample.

12 Head Start is a federal program that provides comprehensive early childhood education, health, nutrition, and parent involvement services to low-income children and families. Five of the ExCEL Quality centers received funding from both Head Start and the Child Care and Development Fund. These "mixed auspice" centers are considered child care settings in this study.

13 Michelle F. Maier, Amy Taub, and Marissa Strassberger, *Variations in Implementation of Quality Interventions: Findings from the VIQI Pilot Study* (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, forthcoming).

14 Cronbach's alphas were calculated for each quality dimension in each two-week reporting period. All three dimensions had good internal consistency, with alphas ranging from 0.73 to 0.95. The dimensions were also distinct from one another, with correlations between interactional quality and structural quality ranging from 0.29 to 0.71, correlations between interactional quality and instructional quality ranging from 0.29 to 0.70, and correlations between structural quality and instructional quality ranging from 0.00 to 0.22.

15 The source of these figures is the ExCEL Quality Teacher Survey.

16 The source of these figures is the ExCEL Quality Parent Information Form. Eighty-six percent of parents consented to their child's participation in the study. However, item response on the parent baseline information form was low. Among the consented sample, approximately a quarter to two-fifths of data was missing for each item. For the analyses, missing data on these covariates were imputed by taking each classroom's mean in the available data. Two centers (consisting of five classrooms) and three additional classrooms were not included in the adjusted analyses because they were missing too much data (such as having no responses on a single variable). Sensitivity tests were conducted where the covariates were imputed using multiple imputation, and the results did not change.

17 This means that there is a 90 percent chance that the true population estimate (that is, the estimate that would be found if this study were conducted in all ECE settings serving low-income families in the United States, rather than just a sample) falls within the range being shown.

18 Cash and Pianta (2014).

19 The interaction between intervention approach and the quadratic time term was statistically significant ($b = 0.001$, $p < 0.05$).

20 This difference was statistically significant ($b = -0.103$, $p < 0.10$).

21 The interaction term was statistically significant ($b = 0.011$, $p < 0.10$).

22 Buell, Han, and Vukelich (2017); Cash and Pianta (2014); Gandhi et al. (2020); Kuger, Kluczniok, Kaplan, and Rossbach (2016); Malmberg et al. (2010).

ACKNOWLEDGMENTS

Many people have contributed to the ExCEL Quality project over the years. The authors would like to thank several MDRC colleagues for their contributions to the preparation of this brief, including Meghan McCormick, Pei Zhu, and John Hutchins for their thoughtful feedback; Isabel Acosta for fact-checking and coordination; Jillian Verrillo for editing the brief; and Carolyn Thomas for preparing it for publication. Importantly, we recognize the contributions of past and present colleagues without whom the ExCEL Quality project would not be possible: from MDRC, Noemi Altman, Rebecca Behrmann, Alexandra Bernardi, Rubin Danberg Biggs, Ilana Blum, Benjamin Bui, Mervett Hefyan, Sharon Huang, Dina Israel, Leah Jacobson, James McCarthy, Seth Muzzy, Ashley Qiang, Sharon Rowser, Amena Sengal, Electra Small, Marissa Strassberger, Mallory Undestad, Evan Weissman, Samantha Wulfsohn, and Jennifer Yeaton; from MEF Associates, Michael Fishman, Carly Morrison, Adelaide Currin, Jan DeCoursey, Emily Ellis, Kim Foley, Marisa Putnam, Lisa Rau, Liza Rodler, and Kate Stepleton; from RTI International, Jean Lennon and Jennifer Keeney; from the University of Colorado at Denver, Julie Sarama, Douglas Clements, Crystal Day-Hess, Laura Ascherl-Morris, Holland Banse, Becky Chance, Candace Joswick, Christina Mulcahy, and Natalie Schock; from Teaching Strategies, Erin Kester and Sandra Faria; and our other affiliated colleague, Sally Atkins-Burnett. Finally, we express deep gratitude to the administrators, coaches, teachers, and children who participated in different phases of the ExCEL Quality project.

This brief was made possible with support from Arnold Ventures as part of the Expanding Children's Early Learning (ExCEL) Initiative.

Dissemination of MDRC publications is supported by the following organizations and individuals that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Arnold Ventures, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JPB Foundation, The Joyce Foundation, The Kresge Foundation, and Sandler Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

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

NEW YORK
200 Vesey Street, 23rd Flr., New York, NY 10281
Tel: 212 532 3200

OAKLAND
475 14th Street, Suite 750, Oakland, CA 94612
Tel: 510 663 6372

WASHINGTON, DC
750 17th Street, NW, Suite 501
Washington, DC 20006

LOS ANGELES
11965 Venice Boulevard, Suite 402
Los Angeles, CA 90066