Variations in Implementation of Quality Interventions

**OMB Information Collection Request
0970 - 0508**

Supporting Statement
Part B

**April 2021
Updated January 2022**

Submitted By:
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

4th Floor, Mary E. Switzer Building
330 C Street, SW
Washington, D.C. 20201

Project Officers:

Ivelisse Martinez-Beck

Amy Madigan

Part B

B1.   Objectives

Study Objectives

The Variations in Implementation of Quality Interventions: Examining the Quality-Child Outcomes Relationship in Child Care and Early Education Project (VIQI Project) is a large-scale, experimental study that aims to inform policymakers, practitioners, and stakeholders about effective ways to support the quality and effectiveness of early care and education (ECE) centers for promoting young children's learning and development. The VIQI Project completed a pilot study in about 40 centers in three metropolitan areas in 2018-2019 that is informing a year-long impact evaluation and process study in 2021-2022 that involves testing the effectiveness of two curricular and professional development models aiming to strengthen the quality of classroom processes and children's outcomes. The objectives of the VIQI project are:

1) Identify dimensions of quality within ECE settings that are key levers for promoting children's outcomes;
2) Inform what levels of quality are necessary to successfully support children's developmental gains;
3) Identify drivers that facilitate and inhibit successful implementation of interventions aimed at strengthening quality; and,
4) Understand how these relations vary across different ECE settings, staff and children.

Generalizability of Results

The study centers will be purposefully selected to produce internally valid estimates of the interventions' causal impact (a key goal of the study), not to promote statistical generalization to other sites or service populations.

The eligible population of centers will include ECE centers that (1) serve low-income children in Head Start and community-based child care settings in metropolitan areas in the United States; (2) provide full-day, center-based early care and education or preschool services; (3) do not already provide substantial amounts of professional development to their teachers; (4) are willing to implement the interventions if selected. The first criterion reflects the target population of centers for the study, and the last three criteria are intended to optimize fidelity of implementation and the interventions' potential to have an impact on classroom quality. The study centers will be located in metro areas to facilitate achieving the sample size targets and the delivery of implementation supports. Metro areas and centers will be purposefully selected for the study based on the feasibility of recruiting the target number of centers and classrooms, while also making sure that the sample is sufficiently diverse to explore research questions focusing on subgroups. In particular, the study team will consider diversity in geographic location (to achieve a broad representation of different regions of the country), setting (to

explore questions related to differences in effects between Head Start centers and community-based child care centers), and levels of classroom quality at the start of the intervention (to examine questions related to variation in effects by initial quality).

Because the study centers will be purposefully selected, the study findings will apply to the centers and children included in the evaluation, which will be located in urban areas and serve high proportions of children from families experiencing poverty. Although not generalizable to the universe of ECE centers in the US, the results are expected to be informative of effects in centers in moderate to large metro areas that serve large proportions of children from families experiencing poverty. The study team will also explore questions related to variation in effects across setting (Head Start and community-based child care), initial quality, and metro area, to try to unpack the extent to which context matters.

Appropriateness of Study Design and Methods for Planned Uses

The strength of this study design and analytical approach is that it will be possible to examine the effects of the two key quality dimensions on children's outcomes, more rigorously and for a more diverse population of ECE centers and children than in prior studies. However, a necessary condition for being able to rigorously and reliably examine the effect of classroom quality on children's outcomes is that the two interventions to which centers are randomly assigned must have a meaningful impact on their focal quality dimension (structural/interactional quality or instructional quality). If this condition is not met, the causal effects of quality on children will be less reliably estimated. To increase the likelihood that the interventions have an impact on quality, the implementation of the two selected interventions by centers will include professional development supports, and the quality dimensions will be measured using instruments with strong validity (e.g. structural validity, concurrent validity) and reliability (e.g. inter-rater reliability, test-retest reliability) in prior studies and that have been shown to be sensitive to intervention effects. The results of this research are intended to inform policymakers, practitioners, and researchers about the effectiveness of the interventions and about which particular quality dimensions and teacher practices may be most important to target to maximize improvements for child outcomes in ECE programming on a large scale. As noted in Supporting Statement A, this information is not intended to be used as the principal basis for public policy decisions and is not expected to meet the threshold of influential or highly influential scientific information.

B2.   Methods and Design

Target Population

The target population includes Head Start and community-based child care centers that serve 3 and 4 year olds from families that predominantly have low incomes, with a goal of recruiting programs that represent a combination of centers that are balanced across Head Start and community-based settings and have varying levels of classroom quality at the start of the intervention.

Sampling and Site Selection

We plan to recruit about 140 centers that serve 3- and 4- year-olds across about 12 metropolitan areas in the United States for the Impact Evaluation and Process Study.

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

As approved under our prior OMB package (OMB #0970-0508), to identify metropolitan areas, the study team gathered information from state and local stakeholders, such as ECE program administrators, local leaders in ECE, and local ECE practitioners, to identify particular metropolitan areas that could be good fits for the VIQI project, using a purposeful, snowball selection strategy. To screen and recruit centers, the study team then reached out to key informants at local administrative entities connected to large numbers of Head Start and community-based child care centers, such as Head Start grantee or delegate agencies or community-based child care programs that operate or oversee multiple child care centers, again using a purposeful, snowball selection strategy.

In May and June of 2021, after obtaining initial screening and eligibility information, the study team will refine and narrow the list of prospective programs and centers and continue outreach to key informants at the program level (and at individual centers when appropriate). Doing so allows us to (1) assess the extent to which the combination of programs and centers on the list of potential candidates provides a distribution of centers that is as balanced as possible across settings (Head Start or community-based child care) which as noted earlier will be examined as a subgroup in the study, and (2) the number of classrooms and children that centers can contribute towards sample size targets. The conversations will occur through a combination of phone calls, video conferences, and in-person meetings (if feasible). In May and June of 2021, we expect to conduct two separate one-hour phone discussions with 100 staff from Head Start centers and community-based child care centers.

From the 140 centers, we assume there will be 3 classrooms per center on average (about 420 classrooms). Within centers, we plan to identify up to one administrator and all lead and assistant teachers in participating classrooms to participate in baseline, follow-up and implementation of fidelity instrument data collection activities. In line with this, we anticipate up to 175 administrators across the participating centers (assumed to be one per center with some additional administrators being added to the group of participants when turnover occurs). We also expect 525 lead teachers and 525 assistant teachers across participating centers (assumed to be one lead teacher and one assistant teacher per classroom with some additional teachers being added to the group of participants when turnover occurs). We expect 59 coaches across the participating centers assigned to one of the intervention conditions (assuming some additional coaches being added to the group of participants when turnover occurs).

For participating classrooms, we also expect to identify and recruit a group of children who are being served to participate in direct assessments. We anticipate asking all parents/guardians of all children in participating classrooms to consent and complete a baseline information form and a report on their child. The information gathered on these forms will be used to identify a list of candidate families and children who are open to participating and meet the selection criteria. We expect 6,300 parents/guardians to be asked a set of baseline information questions and from this we expect to identify about 10 children per classroom to participate in data collection activities (a sample of 4,200 children).

We will aim to generate a group of 3- and 4-year-old children in participating classrooms with sufficient variation in their background characteristics [e.g., their family income (e.g., at or below the federal poverty level and 200% below the federal poverty level), race/ethnicity (e.g., White, Black, Hispanic),

parent's level of education (e.g., at least a high school diploma), dual language learner background (e.g., learning English as a second language)], so that selected group provides sufficient power to detect impacts of the interventions and to explore the relationship of quality to child outcomes for subgroups defined by these characteristics of interest.

B3.   Design of Data Collection Instruments

Development of Data Collection Instruments

Each set of instruments aims to collect unique, but complementary, information about the context and characteristics of centers and programs; experiences, perceptions, and activities of staff (teachers, assistant teachers, coaches, and administrators) in the classrooms and centers; classroom quality; and implementation fidelity. Because limited existing data can inform these constructs of interest in ECE programming, we plan to collect data from multiple sources to enhance our ability to appropriately measure these constructs. See Table B3.1 for information on which project objectives are being addressed by each data collection instrument.

Most instruments were previously approved by OMB (OMB #0970-0508). Through this nonsubstantive change request, we are requesting changes to some of the instruments, including finalizing placeholder items, adjusting some language in instruments to accommodate the use of virtual coaching due to restrictions from the COVID-19 pandemic, confirming the list of child assessments to be collected and adjusting the burden estimate associated with the follow-up protocol for child assessments. See Table B3.2 for a summary of proposed changes to instruments since the last OMB approval.

All survey-based data collection instruments draw from existing scales and measures (e.g., Beliefs about Developmentally Appropriate Practices (DAP) from FACES, Maslach Burnout Inventory, Preschool Teacher's Applied Science Knowledge (PreK-ASK)) whenever possible to capture the range of potential implementation drivers that may facilitate or inhibit successful implementation of the interventions. When an existing item or scale did not exist for the population of interest, the team crafted an item or set of items, with the aim of minimizing the number of questions to be asked of any one respondent type. To minimize measurement error, multiple-items scales were chosen when available.

Coach log items capture dosage and content of coaching sessions provided, as well as the dosage, content, quality and participant responsiveness related to curriculum implementation. Items were developed based on a review of the two curricular models and whatever fidelity-related information was available for the curricula.

Child direct assessments, as well as teacher and parent reports on children, provide standardized and consistent information about children's skills across centers, classrooms and metropolitan areas, since there are no consistent administrative data sources on children's skills that are available. Measures were chosen based on a number of considerations: (1) a need to capture a wide range of children's skills such as math, language, literacy, self-regulation, and executive functioning; (2) whether they have evidence of validity and reliability for 3- to 5-year-old English- and Spanish-speaking children from families with low incomes, (3) whether they have previously shown sensitivity to change from an intervention, and (4) a need to collectively maximize the number of child competencies that could be assessed while staying

within the burden estimates for each timepoint. Final decisions on data collection instruments were made in collaboration with the project's expert consultants.

Table B3.1: Project Objectives Addressed by Each Data Collection Instrument

Project Objective	Instruments
1. Identify dimensions of quality within ECE settings that are key levers for promoting children's outcomes	• Baseline classroom observation protocol • Baseline protocol for child assessments • Baseline teacher reports to questions about children • Baseline parent/guardian reports to questions about children • Follow-up classroom observation protocol • Follow-up protocol for child assessments • Follow-up teacher reports to questions about children • Follow-up parent/guardian reports to questions about children
2. Inform what levels of quality are necessary to successfully support children's developmental gains	• Baseline classroom observation protocol • Baseline protocol for child assessments • Baseline teacher reports to questions about children • Baseline parent/guardian reports to questions about children • Follow-up classroom observation protocol • Follow-up protocol for child assessments • Follow-up teacher reports to questions about children • Follow-up parent/guardian reports to questions about children
3. Identify drivers that facilitate and inhibit successful implementation of interventions aimed at strengthening quality; and,	• Baseline administrator survey • Baseline teacher/assistant teacher survey • Baseline coach survey • Baseline parent/guardian information form • Administrator/teacher COVID-19 supplemental survey questions (incorporated into baseline administrator and teacher surveys) • Follow-up administrator survey • Follow-up teacher/ assistant teacher survey • Follow-up coach survey • Teacher/assistant teacher log • Coach log • Implementation fidelity observation protocol • Interview/ Focus group protocol (administrator, teacher/assistant teacher, and coach burden)
4. Understand how these relations vary across different ECE settings, staff and children	• Baseline administrator survey • Baseline teacher/assistant teacher survey • Baseline coach survey • Baseline classroom observation protocol • Baseline parent/guardian information form • Baseline protocol for child assessments • Baseline teacher reports to questions about children • Administrator/teacher COVID-19 supplemental survey questions (incorporated into baseline administrator and teacher surveys) • Baseline parent/guardian reports to questions about children

Table B3.2: OMB approval status of each data collection instrument, and summary of changes since last OMB approval

Alternative Supporting Statement for Information Collections Designed for Research, Public Health Surveillance, and Program Evaluation Purposes

Instrument	Changes since last OMB approval
Landscaping protocol with stakeholder agencies	• No changes
Screening protocol for phone calls	• No changes
Protocol for follow-up calls/in-person visits for screening and recruitment activities	• No changes
Baseline administrator survey	• No changes
Baseline teacher/ assistant teacher survey	• No changes
Baseline coach survey	• No changes
Baseline classroom observation protocol	• No changes
Baseline parent/ guardian information form	• No changes
Baseline protocol for child assessments	• No changes
Baseline teacher reports to questions about children in classroom	• No changes
Administrator/ teacher COVID-19 supplemental survey questions (incorporated into baseline administrator and teacher surveys)	• No changes
Parent/guardian reports to questions about children	• No changes
Follow-up administrator survey	• No changes
Follow-up teacher/ assistant teacher survey	• Edited wording of some questions about the coaching provided by the project to allow for virtual/remote coaching instead of in-person coaching • Updated the school year to be 2021-2022 For scales that were placeholders, finalized which set of items would be administered Updated the list of languages potentially spoken by teachers to match those spoken by staff in the recruited centers.
Follow-up coach survey	• Updated references to dates • Edited items to accommodate the survey for online distribution (i.e., use the term "select" rather than "choose" when referring to how many response options should be selected, adding references to skip patterns)
Follow-up classroom observation protocol	• No changes
Follow-up protocol for child assessments	• Updated the list of assessments • Edited the reminder letter that will go to parents prior to the spring child assessments to include information about precautions due to the pandemic • Reduced total burden from 3,780 hours to 3,040 due to an increase in the average assessment time (increased from 54 minutes to 57 minutes), a decrease in the number of classrooms (from 420 to 301) and a slight increase in the number of children assessed on average per classroom (from 10 to 10.6)
Follow-up teacher reports to questions about children in classroom	• Removed placeholder items; aligned follow-up version with questions asked in baseline version

	• Added in an item that was accidentally missing from the Internalizing/Withdrawn Items of the Teacher Observation of Child Adaptation-Revised (TOCA-R) scale • As a measure of child engagement, added an item about how often the child attends school. (This item was also included in the baseline teacher report.)
Teacher/ assistant teacher log	• No changes
Coach log	• No changes
Implementation fidelity observation protocol	• No changes
Interview/Focus group protocol	• No changes

B4.   Collection of Data and Quality Control

This section focuses on procedures for data collection activities in the Impact Evaluation and Process Study. The strategies used to collect this information aim to minimize burden and disruption to participants and typical activities in centers.

Data Collected from Screening and Recruitment Instruments (Instruments 1-3)

For the Screening and Recruitment Instruments, regional and local ECE informants, many of whom are expected to be lead staff at Head Start grantees or community-based child care programs that operate or oversee multiple child care centers, will be asked to participate in small-group (or one-on-one) discussions. This format allows the study team to flexibly tailor the questions to the locality or program and ask follow-up questions depending upon gaps in our understanding of ECE programming in a given locality, with the goal of fruitfully and efficiently collecting the needed information to inform the study's screening and sampling criteria. Each facilitator team (pair of two study team members) will make initial e-mail contacts, secure informant participation, and conduct the tailored, semi-structured phone or videoconference discussions (or in-person visits, if feasible).

Facilitator teams will lead the discussion using a subset of the most relevant questions from the semi-structured protocol based on each informant's expertise and our current gaps in knowledge (see Instrument 1: Landscaping Protocol with Stakeholder Agencies and Related Materials, Instrument 2: Screening Protocol for Phone Calls and Related Materials, and Instrument 3: Protocol for In-person Visits for Screening and Recruitment Activities and Related Materials). Each interview will take no longer than 1.5 hours to complete.

Data Collected from Baseline Instruments (Instruments 4-12)

The study team will conduct baseline observations of classroom quality and will ask administrators, lead and assistant teachers, coaches, and parents/guardians of participating children to complete baseline surveys. The study team will also ask a subset of children to complete child assessments.

Baseline surveys. The procedures for collecting the surveys will vary with the study participant, depending on whether direct contact information will be available to the study team and the number of data collection activities the participant will be asked to complete. However, in all cases, participants will receive introductory materials about the study and the purpose of the data collection activity and how

the information being gathered will be handled to maintain their privacy. Contact information for the study team will also be available, so that the participant can ask and have their questions answered, if needed. Study participants will be asked to provide consent or assent prior to completing the surveys. Participants will also be informed that they can refuse to complete the survey, or refuse to answer any of survey questions, and will not be penalized in any way. The surveys will be administered via mixed-mode methodology that consists of online web-based and paper-and-pencil formats. With all approaches, the survey is meant to be self-administered.

Administrators. The study team will collect surveys from administrators in participating centers in Fall 2021. The survey will take 36 minutes to complete. Assent from administrators to complete the baseline survey will be obtained if the participant chooses to complete and return the survey to the study team. The key points covered and information gathered from administrators on the baseline survey are included in Instrument 4: Baseline Administrator Survey. Information regarding communication with administrators (e.g., email) can also be found at the end of Instrument 4.

Lead and assistant teachers. The study team will collect baseline surveys from lead and assistant teachers in Fall 2021. The study team first will provide lead and assistant teachers with an informed consent form that references all data collection activities that the study team will ask them to participate in throughout the course of Impact Evaluation and Process Study. The study team will work with designated study liaisons to distribute the consent forms and baseline surveys in hard copy and/or via email to lead and assistant teachers in the participating classrooms.

If a teacher would like to participate in the survey, s/he will sign (electronically or in hard copy) the consent form, complete the survey, and return both to the study team by mail or electronically. The key points covered and related materials used to contact, consent and gather information from lead and assistant teachers are included in Instrument 5: Baseline Teacher Survey. Information regarding communication with teachers (e.g., letters, email) can also be found at the end of Instrument 5. The survey will take 36 minutes to complete. A $10 honorarium will be provided to centers for each lead and assistant teacher that completes the baseline survey.

Teachers may also be asked to report on about 10 children in their classrooms whose parents/guardians consented to their participation in the study, if COVID-19 circumstances preclude in-person data collection. To the extent possible, children selected for baseline child assessments will be prioritized for teacher reports to questions about children in the classroom. The key points covered and related materials used to gather information from lead teachers are included in Instrument 10: Baseline teacher reports to questions about children in classroom. Information regarding communication with teachers (e.g., letters, email) can also be found at the end of Instrument 10. The reports will take approximately 10 minutes per child to complete. A $4 honorarium will be provided to centers for each child report completed.

Supplemental COVID-19 questions. The study team will also ask administrators and teachers a supplemental set of survey questions that aim to better understand the center's programming due to COVID-19 and to contextualize findings from the Impact Evaluation and Process Study. They will be administered as part of the baseline administrator and teacher surveys. These survey questions will take about 15 minutes to complete. Assent will be obtained if the participant chooses to complete and return

the survey to the study team. The key points covered and information gathered from administrators and teachers are included in Instrument 11: Administrator/teacher COVID-19 supplemental survey questions. Information regarding communication with administrators and teachers (e.g., email) can also be found at the end of Instrument 11.

Coaches. The study team will collect surveys from coaches in Summer/Fall 2021 with some allowance for coaches who are on-boarded late. The survey will take 36 minutes to complete. Assent from coaches is obtained if the participant chooses to complete and return the survey to the study team. The key points covered and information gathered from coaches at the time the baseline survey is administered are included in Instrument 6: Baseline Coach Survey. Information regarding communication with coaches (e.g., email) can also be found at the end of Instrument 6.

Parents/Guardians of children in participating classrooms. Parents or guardians of children being served in classrooms selected to participate in the study will be asked to complete a baseline information form to facilitate identification and selection of children who will be asked to participate in data collection activities for the study.

The study team will provide all parents/guardians with an informed consent form that references the data collection activities that the study team will ask them and their children to participate in throughout the course of the Impact Evaluation and Process Study. The consent form will be available in English and Spanish. The study team will work closely with designated site liaisons to distribute the consent forms and baseline information form to parents/guardians in hard copy and/or via email. If a parent/guardian would like their child to participate in the data collection activities for the study, s/he will sign the consent form (either electronically or in hard copy) and will complete the baseline information form, returning them to the study team. The baseline information form is expected to take 6 minutes to complete. The key points covered and related materials used to contact, consent and gather information from parents/guardians when the baseline information form is administered are included in Instrument 8: Baseline Parent/Guardian Information Form. A $10 token of appreciation will be provided to parents/guardians when asked to complete the baseline information form.

Parents/guardians may also be asked to report on their child's skills, if COVID-19 precludes in-person data collection. Only parents who consent to participate in the study and provide contact information will be invited to participate in the questions on children's skills. The parent report on children's skills will be distributed primarily as a web survey via email, for parents who provide an email address on the baseline information form. Parents may also receive invitations reminders via mailed postcards and phone calls. Parents who do not have internet access may complete the survey by phone. The key points covered and related materials used to gather information from parent/guardians are included in Instrument 12: Parent/guardian reports to questions about children. Information regarding communication with parents/guardians (e.g., letters, email) can also be found at the end of Instrument 12. The reports will take approximately 6 minutes to complete. Parents/guardians will be offered a $10 token of appreciation for completion.

Baseline classroom observations. The study team will aim to conduct observations of classroom quality in all of the participating classrooms at baseline in Fall 2021. Due to the COVID-19 pandemic, these observations may be conducted remotely.

To schedule and conduct these observations, the study team will first contact centers (and any programs overseeing multiple centers) through study liaisons and, if necessary, teachers, to identify potential times that will work for observations during instructional time. At this time, information will also be provided to the liaisons about what is entailed in the observations. Any technology required for remote observations will be provided to the centers, and information on how to use the technology will be provided to center staff and/or teachers. A protocol will guide the introductions and follow-up/wrap up activities with the teachers before and after the observations (See Instrument 7: Baseline Protocol for Classroom Observations.) The protocol will provide teachers with information about the observations and answer any questions they may have (e.g., observation purpose, length of time, privacy, voluntary nature of assessments, OMB statement). This protocol also asks teachers a series of questions about their classroom structure and their practices from that day; these questions will take approximately 18 minutes in total. All observers will be trained on the classroom observation protocol and tool prior to fielding. Observers will also undergo reliability checks for the observational tools collected in line with the recommendations of the tool developers.

Baseline child assessments. Baseline child assessments will be conducted in Fall 2021 after parental/guardian consent has been obtained. The study team will identify and select a subset of 3- and 4-year-old children in each classroom (anticipated to be about 10 children per classroom) whose parents have agreed to allow them to participate in the study and will attempt to stratify children based upon different subgroup characteristics of interest (such as family income [e.g., at or below the federal poverty level and 200% below the federal poverty level], race/ethnicity [e.g., White, Black, Hispanic], parent's level of education [e.g., at least a high school diploma], dual language learner background [e.g., learning English as a second language], and the schedules that children are enrolled in the centers [e.g., child is cared by center for at least 6 hours, 5 days per week]). We will aim to achieve a group with sufficient variation of children with low-income and racially and ethnically diverse backgrounds, so that the selected group of children provides sufficient power to detect impacts of the interventions and to explore the relationship of quality to child outcomes for subgroups defined by characteristics of interest. This sample will constitute the child Impact Evaluation sample. These children will be asked to complete a set of assessments at baseline. Due to the COVID-19 pandemic, these assessments may be conducted remotely.

Children. The study team will schedule and conduct the child assessments by first contacting the centers (and any programs overseeing multiple centers) via designated study liaisons, and if necessary teachers, to provide information about what the assessments entail and identify targeted weeks that will work for the centers and classrooms for conducting the child assessments. The study team will also attempt to identify areas in the centers that can be used to conduct these assessments outside of the classrooms and, if needed, with staff who can aid in the facilitation of the assessments. The study team will plan to conduct the assessments at the potential times identified by the centers and classrooms to minimize disruptions. Prior to conducting the assessments, the assessor will use a protocol that will provide teachers and other staff information about the assessments and answer any questions they may have (e.g., assessment purpose, length of time, privacy, voluntary nature of assessments, PRA statement). The assessor will then ask the teacher or staff member to bring the child to the area being used for assessments and introduce them to the child being assessed. The assessor will make small talk with

child, beginning to build rapport. The assessment battery will take about 30 minutes to complete per child at baseline. The assessments will be offered in English and Spanish. The assessments will be programmed on tablets or laptops, to the extent possible, to facilitate and streamline administration, to reduce errors in administration, and to minimize burden on children. The proposed assessments and related materials used to contact, introduce the assessments, and gather information from center staff, teachers, and children when assessments are administered are included in Instrument 9: Baseline Protocol for Child Assessments.

Due to the young age of the participating children, we will not require signed consent from them to participate. We will have the signed consent of the parents/guardians, and we will collect verbal assent from each child at the start of each assessment period. Should a child not provide assent or wish to stop participating once the assessment has started, they will be returned to their classroom. We will make up to two attempts to assess each child, if they are unwilling to participate or if they are absent on a given day when assessments are scheduled. Upon the completion of the assessments, children will be given stickers to thank them for their participation in the activities.

Data Collected from Follow-Up Instruments (Instruments 13-18)

At follow-up, towards the end of the Impact Evaluation and Process Study, the study team will collect observations of classroom quality and will ask administrators, lead and assistant teachers, coaches, and parents to complete follow-up surveys. The study team will also ask a subset of children to complete direct child assessments. Similar procedures to collecting each data source at baseline will be employed at follow-up. These procedures are detailed below.

Follow-up surveys. The procedures for collecting the surveys will be similar to those used at baseline (see baseline surveys above for more information on procedures).

Administrators. The study team will collect surveys from administrators in participating centers in Spring 2022. If there has been turnover in administrators since Fall of 2021, the new replacement administrators will be targeted for the follow-up survey. The survey will take 30 minutes to complete. Assent from administrators to complete the follow-up survey is obtained if the participant chooses to complete and return the survey to the study team. The key points covered and information gathered from administrators on the follow-up survey are included in Instrument 13: Follow-up Administrator Survey. Information regarding communication with administrators (e.g., email) can also be found at the end of Instrument 13.

Lead and assistant teachers. The study team will collect surveys from lead and assistant teachers in Spring 2022. If there has been turnover in teachers since Fall of 2021, the new replacement teachers will be targeted for the follow-up lead or assistant teacher survey. In addition, lead teachers will be asked to complete a report that includes a set of questions about how children in the child Impact Evaluation sample are doing in the classroom.

The survey will take about 45 minutes to complete. The teacher reports on children will take about 10 minutes per child for teachers to complete, and we will ask teachers to report on about 10 children per classroom. Centers will receive a $15 honorarium for each lead and assistant teacher that completes the follow-up survey. Centers will receive an additional $4 honorarium per child report completed for each

lead teacher. The key points covered and information gathered from lead and assistant teachers at the time the follow-up survey is administered are included in Instrument 14: Follow-up Teacher Survey and Instrument 18: Follow-up Teacher Reports to Questions about Children in Classroom. Information regarding communication with teachers (e.g., letters, email) can also be found at the end of Instruments 14 and 18.

Coaches. The study team will collect surveys from coaches in Spring 2022. If there has been turnover in coaches since Fall of 2021, the new replacement coaches will be targeted for the follow-up survey. The survey will take 30 minutes to complete. The key points covered and information gathered from coaches are included in Instrument 15: Follow-up Coach Survey. Information regarding communication with coaches (e.g., email) can also be found at the end of Instrument 15.

Parents/Guardians. In Spring 2022, parents or guardians of children being served in participating classrooms may be asked to report on their child's skills, if COVID-19 circumstances preclude in-person data collection. The reports on children will take about 6 minutes to complete. Parents/guardians will be offered a $10 token of appreciation for completion. The key points covered and information gathered are included in Instrument 12: Parent/Guardian Reports to Questions about Children in Participating Classrooms. Information regarding communication with parents/guardians (e.g., letters, email) can also be found at the end of Instrument 12.

Follow-up classroom observations. Targeting the same classrooms participating at baseline, the study team will aim to conduct up to three observation timepoints of classroom quality in all of the participating classrooms at follow-up in the Winter/Spring 2022. The observer will also ask teachers a series of questions about their classroom structure and their practices from that day; these questions will take approximately 18 minutes in total. See baseline classroom observations above and Instrument 16: Follow-up Classroom Observation Protocol for more information on procedures.

Follow-up child assessments. Targeting the same children participating at baseline, the study team will aim to complete a set of direct assessments in Spring 2022. The assessment battery will take about 57 minutes to complete, on average, per child at follow-up. We anticipate that approximately 18% of the children assessed will speak Spanish at home and, therefore, will need to be assessed on the language assessment in both English and Spanish; the assessment battery will take approximately 70 minutes for these children. For the rest of the child sample, we expect the assessment battery to last for approximately 54 minutes. See baseline child assessments above and Instrument 17: Follow-up Protocol for Child Assessments for more information on procedures.

Data Collected from Implementation Fidelity Instruments (Instruments 19-22)

The implementation fidelity instruments will be collected throughout the Impact Evaluation and Process Study. The procedures for collecting the information vary, depending upon the data source.

Teacher logs. Beginning in September 2021 and ending in June 2022, the study will ask all lead and assistant teachers in participating classrooms across research conditions to complete weekly logs. If turnover in teachers occurs, the new replacement teachers will be asked to complete the logs.

The logs will be available online, and teachers will be trained to log onto and use a data system to complete the logs. Each log is expected to take about 15 minutes to complete. Email and/or text message notifications will be sent to teachers to remind them to complete the logs. Information about who to contact for questions or to address technical issues in completing the logs will also be provided to teachers. The key points covered and related materials used to introduce the logs with teachers, how to complete the logs, how the information will be used, and how the information will be protected are included in Instrument 19: Teacher Log. Information regarding communication with teachers (e.g., email, text messages) can also be found at the end of Instrument 19. Centers will receive an additional $10 honorarium per month for each lead and assistant teacher completing the teacher logs.

Coach Logs. Beginning in September 2021 and ending in June 2022, we will ask coaches hired to support the installation of one of the interventions to complete logs after each coaching session with teachers in participating centers and classrooms. If turnover in coaches occurs, the new replacement coaches will be asked to complete the logs. Information gathered in the logs will support coaches' management of their caseloads. We expect that coaches will need to track and monitor this information, regardless of whether the process study was ongoing. This information will be shared with the research team to track and monitor the delivery of professional development and implementation of the interventions as well.

The logs will be available online, and coaches will be trained on how to access a data system to complete the logs during the onboarding process. Each log is expected to take about 15 minutes to complete. The coach logs will be administered using the approach described above for the teacher logs. The key points covered and related materials used to introduce the logs to coaches, how to complete the logs, how the information will be used, and how the information will be protected are included in Instrument 20: Coach Log. Information regarding communication with coaches (e.g., email, text messages) can also be found at the end of Instrument 20.

Implementation fidelity observations. The study team will aim to conduct observations to assess fidelity of implementation of the interventions in a subset of classrooms assigned to each of the intervention conditions. We will also conduct these observations in a subsample of classrooms that are assigned to the control condition to assess the extent to which specific behaviors, practices and activities supported by the interventions are evident in the control classrooms to inform the potential relative treatment contrast across research conditions. The observations will consist of one time-point of observations conducted in Winter/Spring 2022. Depending on circumstances surrounding COVID-19, these observations may be remote.

To identify classrooms to participate in these observations, the research team will select a subset of centers that are stratified by whether they provide Head Start or community-based child care services and high or low classroom quality at baseline. Within these centers, the study team will select one classroom at random to participate in the implementation fidelity observation (for up to 80 classrooms).

To schedule and conduct these observations, the study team will first contact centers (and any programs overseeing multiple centers) via study liaisons, and, if necessary teachers, to provide information about what the observations entail and identify potential times that will work for the centers and classrooms to conduct the observations during instructional time. Any technology required for remote observations will be provided to the centers, and information on how to use the technology will be provided to center

staff and/or teachers. A protocol will be used to guide the introductions and follow-up/wrap up activities with the teachers before and after the observations. Upon arriving at the centers, the member of the study team or observer will use this protocol that will provide teachers information about the observations and answer any questions they may have (e.g., observation purpose, length of time, privacy, voluntary nature of assessments, PRA statement). This protocol also asks teachers a series of questions about their classroom structure and their practices from that day; these questions will take approximately 18 minutes per observation. The key points covered and related materials used to contact, gather information from centers and to introduce the observations to center staff and teachers, and to guide the pre- and post-observation discussions are included in Instrument 21: Implementation Fidelity Observation Protocol.

Interviews/focus groups. The study team will conduct qualitative interviews with a subset of participating administrators, coaches, and teachers in Winter 2022. A random subset of administrators (up to 8 administrators within 4 localities) and coaches (up to 3 coaches within 4 localities) will be asked to participate in a one-on-one interview. A random subset of lead and assistant teachers (up to 48 teachers within 4 localities) will be asked to participate in a small-group/one-on-one interview where teachers across centers are interviewed in separate groups by position.

Each one-on-one interview or small-group interview will last up to 1.5 hours. Interviews aim to gain insights from study participants on their experiences implementing the interventions, engaging in professional development and completing the data collection instruments. The interviews will be facilitated and led by a member of the research team using a semi-structured protocol that will be adapted depending upon the participants being interviewed. The one-on-one interviews are expected to be conducted by phone and the small-group interviews are expected to be conducted by video conference or in person, depending on the circumstances surrounding the COVID-19 pandemic.

To identify interview participants, the research team will select a random subset of centers (up to 8 centers within 4 localities) that are stratified by whether they provide Head Start or community-based child care services and high and low quality at baseline. Within these centers, the study team will look to interview staff at the different levels. The study team will contact the coaches and administrators directly to ask if they would be willing to participate in a one-on-one interview of their experiences in the study. The study team will work with the administrators and coaches to identify times for the individual interviews. Administrators will also be notified that lead and assistant teachers will be contacted separately. The study team will then contact lead and assistant teachers directly to ask if they would be interested and willing to participate in a small group interview. The study team will propose times for the small group interviews with lead or assistant teachers, and those who are available and interested will confirm whether the proposed times work for them. Information will also be provided to administrators, coaches, and teachers about the purpose of the interviews and what information will be gathered, how the information will be used, how the study team will protect their information, the voluntary nature of the data collection activity, and who to contact should they have questions about the data collection activity (e.g., interview purpose, length of time, PRA statement). The key points covered and related materials used to introduce the interviews, how the information will be used, and

how the information be handled to maintain the privacy are included in Instrument 22: Interview/Focus Group Protocol.

B5.   Response Rates and Potential Nonresponse Bias

Response Rates

The expected response rates vary by instrument, time point, and participant type. The response rates for each instrument are expected to be similar to what was achieved in the VIQI pilot study.

Screening and Recruitment Instruments

For screening and recruitment materials, maximum response rates are critical to ensuring that the study team selects the most appropriate centers that meet the sampling criteria for participating in the VIQI project. We anticipate that the vast majority of informants will likely be interested in providing their insights to help inform the screening and recruitment of metropolitan areas and centers. As such, we expect little nonresponse. Using the study team's past experiences with engaging similar informants to collect information for screening and recruitment purposes, the team expects approximately <u>80 percent of targeted participants to respond to each protocol for screening and recruitment</u>.

Baseline Instruments

We expect nearly <u>100 percent to respond to the baseline surveys for administrators and coaches</u>. In the VIQI pilot, 98% of administrators and 100% of coaches completed the baseline survey. We expect about <u>85 percent of lead and assistant teachers to consent to participating in the study and to complete baseline surveys/reports on children</u> in participating classrooms. In the VIQI pilot study, 88% of lead teachers completed a baseline survey.

To collect consent and baseline information form from parents/guardians of children being served in participating classrooms, we will target almost all parents/guardians of children in the classrooms and we expect that <u>85 percent of parents/guardians of children</u> in participating classrooms will return the consent forms on behalf of their children and will complete the baseline parent/guardian information form and report on questions about their child that accompany the consent form. In the VIQI pilot study, response rates were lower than expected (45% of parents returned a signed consent form and 42% completed the baseline information form). However, in a similar study, called the ExCEL Quality study, which examined a second year of implementation within many of the same localities and centers participating in the VIQI pilot, 90% of parents returned a signed consent form and 68% completed the baseline information form.

We will aim to collect two time-points of classroom observations in all participating classrooms. We expect about <u>80 percent completion of the baseline observations</u> in participating classrooms given the need for remote observations due to the COVID-19 pandemic. In the VIQI pilot study, 100 percent of classrooms have complete baseline observations, but these were in-person observations.

We will aim to collect baseline child assessments for a selected sample of children in participating classrooms. We expect <u>85 percent of selected children to complete the baseline child assessments</u>.

Follow-up Instruments

At follow-up, we expect similarly high response rates. We expect nearly <u>100 percent to respond to the follow-up surveys for administrators and coaches</u>. In the VIQI pilot, 76% of administrators and 92% of coaches responded to the follow-up survey. We expect about <u>95 percent of lead and assistant teachers who consented and completed the baseline survey to complete a follow-up survey and reports on children</u> in participating classrooms. In the VIQI pilot, 94% of lead teachers responded to the follow-up survey.

We will aim to collect three time-points of follow-up classroom observations in all participating classrooms. We expect <u>nearly 100 percent completion of the follow-up observations</u> in participating classrooms. In the VIQI pilot, 100 percent of classrooms have complete follow-up observations.

We will aim to collect follow-up direct child assessments for a selected subset of children in participating classrooms. We expect <u>85 percent of selected children to complete the follow-up child assessments</u>.

Implementation Fidelity Instruments

Throughout the Impact Evaluation and Process Study, we will aim to collect logs from teachers and coaches supporting the installation of the interventions and teachers across research conditions. We expect that <u>80 percent of lead and assistant teachers will respond to at least one log</u>, with about <u>50 percent of the total number of logs completed</u>. We expect nearly <u>100 percent of coaches to respond to at least one log</u> throughout each phase of the study, with about <u>90 percent of the total number of logs expected to be completed</u>. In the VIQI pilot, 85% of the total number of expected coach logs were completed. In the VIQI pilot, teacher logs were not used.

As part of the Process Study, a subset of centers and their underlying classrooms will be selected to participate in implementation fidelity observations. We expect nearly <u>100 percent of the subset of classrooms will participate in the implementation fidelity observations</u> across research conditions. In the VIQI pilot, over 90 percent of the subset of classrooms targeted for fidelity visits participated in implementation fidelity observations.

The interviews/focus groups are not designed to produce statistically generalizable findings and participation is wholly at the respondent's discretion. Response rates will not be calculated or reported.

Maximizing Response Rates

To ensure that the VIQI project has sufficient power to address the research questions of interest for different phases of the project, it will be important to reach the expected response rates described above. We fully recognize potential challenges and have structured a data collection plan accordingly. Our plan draws upon our extensive experience managing and collecting similar sets of data from children, teachers, classrooms, coaches, and ECE centers in multiple large-scale, longitudinal, and experimental studies, and includes the following strategies:

- **Building relationships.** Our research team is comprised of seasoned operations staff across MDRC and MEF Associates who have worked extensively with ECE centers across the United States in large-scale studies not only to maintain strong relationships and work collaboratively with centers but also to trouble shoot and provide technical assistance when necessary to

minimize disruptions and facilitate data collection activities. We will also work with Head Start grantees and programs that have oversight over multiple child care centers (and individual centers to the extent necessary) that designate liaisons who will coordinate with the study team to facilitate data collection activities.

- **Minimizing burden.** We also draw upon our expertise and experience to put in place mixed-mode administration for the instruments whenever possible to minimize burden on study participants. Further, the instruments and protocols will be developed to be streamlined, cleanly formatted, and as brief as possible. We will draw upon principles from behavioral economics to tailor contact and communication with study participants to encourage responses. We will also aim to balance the breadth of data being collected by minimizing burden and disruptions to centers, staff and children by optimizing the amount of data collected at each observation or assessment point. Last, we will be flexible in accommodating the schedules of centers and classrooms when collecting data, while still adhering to the planned timeline for data collection activities.

- **Conversion and avoidance of refusals.** We will train fielding staff of the instruments in conversion and avoidance of refusals, including training on distinguishing "soft" refusals from "hard" ones. Soft refusals often occur when a study participant has been reached at an inopportune time. In these cases, it is important to back off gracefully and to establish a convenient time to follow up with the study participant, rather than to persist at the moment. Hard refusals do occur and must also be accepted gracefully by the fielding staff.

- **Oversight.** Our team also includes Abt/SRBI and RTI who will lead data collection efforts in participating centers and have extensive experience collecting high-quality classroom- teacher- and child-level data in large-scale studies. Abt/SRBI's senior data collection manager will provide centralized oversight of the collection of lead and assistant teacher consents, surveys, and reports on children and classroom observations, while RTI will provide centralized oversight to the collection of parent/guardian consents, baseline information forms, reports on children, and child assessments. Staff experienced in managing early childhood data collection efforts will be hired as field supervisors to oversee data collection efforts at each locality. Field supervisors will hire and train local field staff to conduct each data collection activity. Abt/SRBI and RTI will design, implement, maintain, and document an integrated study database that will provide oversight of all data collection activities. Such a system is critical for allowing project staff to monitor the flow of information and ensure that each designated sample unit (child, parent, teacher, coach, administrator, etc.) is properly surveyed and that all required information is obtained, identified, and stored. Further, MDRC will have a dedicated data collection coordinator who will work closely with both survey firms and will have oversight over all of their data collection activities. MDRC, leveraging the operational and TA/monitoring activities of MEF/MDRC operational team members, will also have direct oversight over the collection of administrator surveys, teacher and coach logs, and coach surveys. Thus, between Abt/SRBI and RTI staff and MDRC/MEF operational staff, our team will be in contact with centers at regular

18

intervals, allowing us to follow up with centers and respondents on a frequent basis to ensure high response rates.

- **Monitoring.** The study team will closely monitor data collection and response rates by data source to ensure high response rates and no differential response rates by research conditions. Weekly meetings will address any issues that arise during preparations for data collection and data collection itself. The team will produce internal monthly progress reports, which will include any issues and solutions for correcting issues. The study team will also review early files of data collected from each instrument to assess if there are any issues in the completeness or quality of the data being collected, so that issues can be quickly identified and solved early in the fielding stages of each instrument.

NonResponse

Although participants will not be randomly sampled and findings are not intended to be representative of a population, it is important to assess non-response bias to ensure that the respondent sample reflects a representative group of the randomly assigned participants. Two types of bias will be assessed: (1) differences in response rates and respondents' characteristics *across the experimental groups* in the design (differential response) and (2) differences in the characteristics of *respondents compared to non-respondents*. The first type of bias affects whether the impacts of the interventions are confounded with pre-existing differences between experimental group and control group respondents (internal validity), while the second type of bias affects whether the results from the study can be generalized to the wider group of eligible respondents (external validity).

Several tests will be conducted to assess whether differential non-response is compromising the internal validity of the experimental design. For each data source:

- Response rates by experimental group will be compared to make sure the response rate is not significantly higher for one research group.

- A multinomial logistic regression will be conducted among respondents. The "left hand side" variable will be random assignment group membership while the explanatory variables will include a range of baseline characteristics. An omnibus test such as a log-likelihood test will be used to test the hypothesis that the set of baseline characteristics are not significantly related to a respondent's experimental group. Failure to reject this null hypothesis will provide evidence that respondents are similar across experimental groups.

The guidelines provided by the What Works Clearinghouse at the Institute of Education Sciences (Department of Education) will be used to determine whether attrition is "low" or "high" based on these analyses. If these tests indicate that differential non-response is "high," we will regression-adjust the impact analyses using respondents' baseline characteristics and outcomes. To make sure that the regression-adjustment is adequately removing the bias, we will conduct a sensitivity test where we will drop the random assignment blocks where differential response rates are the largest, and then estimate impacts based on this smaller sample.

To examine whether the results are generalizable to the eligible population (externally valid), the following analysis will be conducted for each data source:

- The baseline characteristics of respondents will be compared to the baseline characteristics of non-respondents using a logistic regression where the outcome variable is whether someone is a respondent and the explanatory variables are baseline characteristics. An omnibus test such as a log-likelihood test will be used to test the hypothesis that the set of baseline characteristics are not significantly related to being a respondent. Failure to reject this null hypothesis will provide evidence that non-respondents and respondents are similar at baseline.

If these tests indicate that respondents are different from non-respondents, the presentation of the findings will clarify that the results may not be generalizable to the full group of eligible respondents. As a sensitivity analysis, we will also reweight the respondent groups to reflect the characteristics of the full group of eligible participants, to explore whether the results could differ.

For the impact analyses, baseline data will be used as covariates in the analysis to describe the respondents and improve precision. Therefore, it will be acceptable to impute these baseline variables using an appropriate method such as multiple imputation. Follow-up data will not be imputed.

B6. Production of Estimates and Projections

The research questions for the Impact Evaluation will be answered using a 3-group random assignment research design. Centers will be randomly assigned to one of three groups: a group that receives Intervention A (Group 1), a group that receives Intervention B (Group 2), or a group that continues to conduct "business as usual" (Control) in approximately equal proportions (see table B6.1). Each of the two selected interventions will target a different dimension of quality—one will target structural/interactional quality, and the other will target instructional quality. If the interventions improve quality as intended, this design will create random (experimentally-induced) variation in the two quality dimensions that can be used to rigorously estimate their effect on children's outcomes using an instrumental variables (IV) analysis. The findings will represent intervention effects and quality effects for the population of centers in the study.

Table B6.1 Number of Centers Assigned to Each Treatment Group

Experimental Condition	Centers
Business as Usual Control	46
C4L	47
Creative Curriculum	47

Prior to the impact analyses, we will assess whether the three experimental groups in the study are similar. Because the baseline quality and child outcome measures will be collected *after* the implementation of the interventions has started, they are not a true "baseline" and there is a possibility that we may see baseline differences across groups (early impacts). Therefore,

we will assess whether random assignment produced equivalent groups by comparing characteristics that are unlikely to be impacted by the interventions, such as student and teacher demographics.

The analysis of the Impact Evaluation data will examine different types of effects: (1) the effect of each intervention on classroom quality and teacher and child outcomes; (2) the combined intervention (that is, the effect of both interventions pooled together) on classroom quality and child outcomes; (3) the effect of each targeted dimension of quality (structural/interactional quality and instructional quality) on child outcomes; and (4) the effect of global quality (i.e., a composite measure of the two quality dimensions) on child outcomes.

The estimation and analysis for each type of effect is described below. In drawing inferences about these estimated effects, standard statistical tests such as t-tests (for continuous variables and dichotomous measures) or chi-square tests (for categorical measures) will be used to account for estimation error and determine whether estimated effects are statistically significant. Each of the analyses will be conducted for the full group of participating centers, as well as for subgroups of interest [e.g., by centers' initial levels of quality at baseline (low vs. high) and by setting (Head Start vs. community-based)]. Subgroup analyses will be conducted either by running separate models for each subgroup of interest, or by adding subgroup interactions in the models. Our sampling design does not require the use of survey or sampling weights. Before the analysis begins, we will pre-register the analysis plan with Open Science Framework.

Once analysis for the VIQI impact evaluation and process study is completed and the results are published, we plan to archive the analysis variables and measures with documentation to support accurate estimates and projections for secondary analysis. This documentation will include code books, user manuals, file structure, variables, sample weights (if applicable), and methods. The documentation will also include information about the types of data collected, data handling procedures and storage methods, procedures for data prep and level of restriction for different data.

B7. Data Handling and Analysis

Data Handling

Electronic notes taken during screening and recruitment activities, as well as interviews/focus groups, will be stored in a secure, password-protected location. Any audio recordings from interviews/focus groups will be gathered on secure, password-protected audio recorders or done via Zoom recordings that are housed in the secure Zoom cloud before the research team transfers them at the end of the day to a secure, password-protected Amazon Workspaces government-cloud project folder that only the study team has access to.

All web-based surveys will be programmed and tested prior to fielding to ensure accurate administration and minimize errors during data processing. When permissible, we will program instruments, eliminating one step in digitization and allowing immediate corrections and eliminating the number of steps where error could be introduced. We will also program in validity and consistency checks. For some paper forms, scanning will be appropriate. Whether scanning or conducting manual data entry,

quality assurance protocols will be put in place to ensure accurate recording of data. Questionable values will be checked against source documents.

Data Analysis

Analysis of Impact Study Data. The statistical approach for the impact analysis is different for each research question. The analysis for each question is described in detail below.

Effect of each intervention on classroom quality and teacher and child outcomes. The effect of each intervention will be estimated by comparing the classroom quality and teacher and child outcomes of centers assigned to each intervention group (Group 1 or Group 2) to those of centers assigned to the control group (intent to treat). In practice, these analyses will be conducted using a model that regresses the outcome of interest (classroom quality, teacher outcome, or child outcome measure) against indicators of group membership (Group 1 and Group 2). The regression coefficient on these indicators will provide an intent-to-treat estimate of the effect of each intervention on the outcome of interest. The model will also include a set of random assignment block indicators, to account for the random assignment design and to improve the precision of estimated effects. Because random assignment occurs at the center level, that is the highest level of clustering that needs to be accounted for in the analysis. Higher levels of potential clustering, such as locality and program, will be accounted for as covariates in the model. In addition, the model will control for measures of classroom-level (such as classroom composition and baseline quality) and child-level baseline characteristics and baseline outcomes; because of random assignment, controlling for these baseline characteristics and outcomes in the model is not strictly necessary and it will not affect the impact estimates, but we will include them to improve the *precision* of estimated effects (reduce their standard error). The analysis will use a multi-level modelling structure to account for the clustered nature of the data: a two-level model will be used for classroom quality (classrooms nested in centers) and a three-level model will be used for child outcomes (children nested in classrooms nested in centers).

Combined intervention effect on classroom quality and child outcomes. The combined effect (intent to treat) of the two interventions will be estimated by comparing the quality and child outcomes of centers in Group 1 and 2 to those of centers assigned to the control group. The statistical model will be similar to the one previously described, except that the key independent variable will be an indicator of assignment to Group 1 or Group 2.

Effect of each targeted quality dimension on child outcomes. The effect of structural/interactional quality and of instructional quality on children's outcomes will be examined using an IV approach. In practice, the effect of these two quality dimensions will be estimated using two IV-based methods. The first method will be a two-stage least squares (2SLS) analysis, where indicators of group membership (Group 1 and 2) will be used as the instruments. In the first stage models, each quality dimension (structural/interactional quality and instructional quality) will be regressed against the two instruments (an indicator of assignment to Group 1 and an indicator of assignment to Group 2), a set of random assignment block indicators, and classroom-level and child-level baseline characteristics. From these regressions, the predicted values of the two quality dimensions will be obtained. These predicted values represent variation in the two quality dimensions that is experimentally induced. In the second stage model, the child outcome of interest will be regressed against the two predicted quality dimensions, as

well as random assignment block indicators, and classroom and child-level baseline characteristics. The regression coefficients on the predicted quality dimensions will provide estimates of the effect of the two quality dimensions on children. These estimates are unbiased if these two dimensions are the only pathways through which the interventions improve children's outcomes. The analysis will use 3-level models to account for the clustered nature of the data (children nested in classrooms nested in centers). To explore whether the effect of the two quality dimensions on child outcomes is *non-linear*, we will estimate the effect of each quality dimension for subgroups of centers defined by their baseline quality (low versus high). If effects are larger for one group compared to the other, this would suggest that the effect of a given dimension of quality may be non-linear. This analysis is non-experimental and will be considered more exploratory.

The second method will be to use a multi-site multi-mediator IV approach (Bloom et. al, 2000). This approach makes it possible to relax the assumption that the quality dimensions are the only pathway through which quality affects children, but the effect of each quality dimension must be estimated separately (one at a time). The effect of a quality dimension will be estimated by regressing the impact of each intervention on the child outcome of interest, against the impact of each intervention on the quality dimension of interest. The slope of the regression line represents the estimated effect of the quality dimension of interest on the child outcome (this slope can also be calculated manually as the differential effect between the two interventions on the child outcome of interest, divided by the differential effect between the two interventions on the quality dimension of interest). The intercept of the regression line represents the effect of all other (unmeasured) quality dimensions. The standard error of the effect of a quality dimension on children is equal to the standard error of the differential effect of the interventions on the child outcome divided by the differential effect on the quality dimension. This standard error can be used for hypothesis testing. If there is variation in the effect of a quality dimension across auspice (Head Start, Child Care), by initial quality (low versus high), or another baseline center characteristic, then this variation will be used to create additional instruments that will be used to explore whether the shape of the relationship between global quality and children's outcomes may be nonlinear. This analysis is non-experimental and will be considered exploratory.

Effect of global quality on child outcomes. The effect of global quality (a composite measure of the two dimensions of quality) on children's outcomes will also be estimated using an IV approach, using the two methods described above. For the first method (two-stage least squares), in the first stage model, global quality will be regressed against the instrument (an indicator for whether a center was assigned to Group 1 or Group 2), a set of random assignment block indicators, and classroom-level and child-level baseline characteristics. From these regressions, predicted values of global quality will be obtained. In the second stage model, the child outcome of interest will be regressed against predicted global quality, as well as random assignment block indicators, and classroom and child-level baseline characteristics. The regression coefficient on predicted global quality will provide an estimate of global quality on children. This estimate is unbiased if global quality is the only pathway through which the interventions improve children's outcomes. If one of the interventions has a larger effect on global quality than the other, then we will be able to rigorously examine whether the effect of global quality is nonlinear, by using treatment group membership (to Intervention 1 and to Intervention 2) as instrumental variables for global quality and its quadratic (two mediators).

For the second method (mutli-site multi-mediator IV), the effect of global quality (and its nonlinearity) will be estimated using the same strategy described earlier, except that the quality measure of interest will be global quality (instead of each quality dimension). As noted, this approach does not assume that global quality is the only pathway through which the interventions affect children's outcomes.

Baseline analyses. Prior to conducting the impact analyses, we will compare the baseline characteristics and outcomes of centers, teachers/classrooms, and children in the three experimental groups, to confirm that random assignment has produced three groups that are similar at baseline. We will test whether the 3 groups' characteristics are statistically different from each other for each baseline characteristic and systematically across all characteristics.

Analysis of Process Study Data. A variety of descriptive and comparative techniques will be used to describe implementation drivers and mean levels and variation in different dimensions (dosage, adherence, quality) of fidelity of the intervention (including the provision of professional development and delivery of the curriculum). Correlational analysis will be used to examine associations among various implementation drivers as well as among implementation drivers and aspects of fidelity. Additionally, the achieved relative strength (that is, the treatment contrast) between treatment and control classrooms will be calculated based on procedures detailed by Hulleman and Cordray (2009) (standardizing the average difference between fidelity indices from each condition). Achieved relative strength will be used to help interpret the findings from the impact analysis, and many of the constructs created as part of the process study can be used as moderators (e.g., center readiness) in the impact analysis. Analysis will be conducted with all centers and separately for subgroups with data on outcomes of interest. Our analysis will take into account the nested nature of the data as needed. We will consider methods for handling missing data (as discussed above). We will employ data reduction techniques and psychometric work (e.g., internal consistency, factor analysis, concurrent validity), particularly for fidelity of implementation variables that are collected across time and for new measures. Any qualitative data collected during center visits may also be quantified to measure specific readiness factors of interest (e.g. "rating" a site's level of quality assurance and improvement processes as minimal, moderate, or strong).

One notable contribution of this process study is an ability to examine center "readiness" and which factors individually, and in combination, predict different aspects of fidelity of implementation. Profile analysis may be used to identify whether centers can be grouped based on their initial "levels" on different readiness factors. Identifying profiles of readiness can inform who may be ready to take on a new initiative and who may need more support before doing so; and for gaining further insight into why (or why not) an intervention was effective at changing quality.

Finally, because implementation drivers, as well as implementation efforts and their outcomes, may naturally change over time and/or may be affected by intervention-induced changes in quality, we plan on examining changes in these constructs. This will allow us to see whether there are differences in a variety of implementation-related factors—such as beliefs, center readiness for change, and organizational climate—before and after implementation of a quality improvement intervention.

Data Use

Reports and/or briefs will be published to summarize the findings from the Impact Evaluation and Process Study. It will include details on the data analyses conducted, interpretation of the findings, and study limitations. Findings aim to answer key open questions in the field about how different dimensions of classroom quality are related to a range of child outcomes and whether there are thresholds in these associations.

B8. Contact Person(s)

JoAnn Hsueh, MDRC, joann.hsueh@mdrc.org
Michelle Maier, MDRC, michelle.maier@mdrc.org
Marie-Andree Somers, MDRC, marie-andree.somers@mdrc.org
Noemi Altman, MDRC, Noemi.Altman@mdrc.org
Electra Small, MDRC, Electra.Small@mdrc.org
Margaret Burchinal, University of Virginia, kqu4rg@virginia.edu
Jean Lennon, RTI International, jlennon@rti.org
Jennifer Kenney, RTI International, jkeeney@rti.org
Ricki Jarmon, Abt Associates, ricki_jarmon@abtassoc.com
Brenda Rodriguez, Abt Associates, Brenda_Rodriguez@abtassoc.com

Attachments

Appendices
Appendix A: Conceptual Framework
Appendix B: Certificate of Confidentiality

Instruments
Instrument 1: Landscaping Protocol with Stakeholder Agencies and Related Materials
Instrument 2: Screening Protocol for Phone Calls and Related Materials
Instrument 3: Protocol for In-person Visits for Screening and Recruitment Activities and Related Materials
Instrument 4: Baseline Administrator Survey
Instrument 5: Baseline Teacher Survey
Instrument 6: Baseline Coach Survey
Instrument 7: Baseline Protocol for Classroom Observations
Instrument 8: Baseline Parent/Guardian Information Form
Instrument 9: Baseline Protocol for Child Assessments
Instrument 10: Teacher reports to questions about children in classroom
Instrument 11: Administrator/teacher COVID-19 supplemental survey questions
Instrument 12: Parent/guardian reports to questions about children
Instrument 13: Follow-up Administrator Survey
Instrument 14: Follow-up Teacher Survey
Instrument 15: Follow-up Coach Survey
Instrument 16: Follow-up Classroom Observation Protocol
Instrument 17: Follow-up Protocol for Child Assessments

Instrument 18: Follow-up Teacher Reports to Questions about Children in Classroom

Instrument 19: Teacher Log

Instrument 20: Coach Log

Instrument 21: Implementation Fidelity Observation Protocol

Instrument 22: Interview/Focus Group Protocol