*AERA Open*
*January-December 2021, Vol. 7, No. 1, pp. 1–21*
*DOI: 10.1177/23328584211044519*
*Article reuse guidelines: sagepub.com/journals-permissions*
*© The Author(s) 2021. https://journals.sagepub.com/home/ero*

# Linking Features of Structural and Process Quality Across the Landscape of Early Education and Care

**Emily C. Hanno**
*Harvard University*

**Kathryn E. Gonzalez**
*Harvard University*
*Mathematica Policy Research*

**Stephanie M. Jones**
**Nonie K. Lesaux**
*Harvard University*

*Commonly regulated structural quality features, like educator education levels and group size, are thought to be foundational to the quality of children's everyday experiences in early education and care settings. Yet little is known about how these features relate to the day-to-day interactions and activities that occur in these settings—or process quality features—across the landscape of early education and care. In this study, we examine the association between structural quality features and process quality features in a diverse sample of classrooms (n = 672) participating in a statewide study of early education and care. Using a permutation test approach, we found that group size and child-to-adult ratio were most consistently linked to children's experiences but educator education, experience, and curriculum usage were largely unrelated. Implications of these findings for quality improvement initiatives are discussed.*

Keywords: *process quality, structural quality, early education and care*

High-quality early education and care programs are seen as a key policy approach for preparing children for kindergarten and beyond. Yet few programs today meet the quality standards considered necessary for supporting young children's development (Chaudry et al., 2017). In response, the number of systemwide quality improvement initiatives—or Quality Rating and Improvement Systems (QRIS)—has nearly tripled in the United States between 2007 and 2017 (Build Initiative & Child Trends, 2019). Despite a great deal of investment in such initiatives, we still have little evidence on the critical, regulable features of early education and care programs that foster high-quality, developmentally supportive experiences for young children in these settings. This article addresses this broad issue by applying a novel empirical approach in a large, statewide data set to examine whether certain structural quality features consistently predict the everyday interactions and activities in early education and care settings.

Structural quality features include the readily quantifiable and regulable features of settings, like staff qualifications, group size, and classroom materials, hypothesized to underlie process quality features (Phillipsen et al., 1997). Process quality has been variously defined in the literature but generally reflects what actually happens on a day-to-day and moment-to-moment basis in settings, including activities and interactions between and among adults and children (Cryer et al., 1999; Howes et al., 2008; Pianta et al., 2005). Process features are considered the most proximal determinants of children's learning and have been proposed as a key mechanism linking structural features to children's outcomes (Hamre, 2014; Markowitz et al., 2017). Specifically, structural features are theorized to set the stage for process quality by creating the conditions in which high-quality, developmentally supportive processes can occur (NICHD Early Child Care Research Network [ECCRN], 2002; Slot, 2018; Slot et al., 2015). This hypothesis largely motivates the widespread regulation of structural features: In 2016, 100% of QRIS included stipulations for minimum educator qualifications, approximately 80% set curricular standards, and more than 50% considered group size and adult-to-child ratios as quality indicators (National Center for Early

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Childhood Quality Assurance, 2017). Yet there exists mixed empirical evidence in support of the hypothesis that structural features drive process quality (e.g., Burchinal et al., 2002; Cryer et al., 1999; Hanno et al., 2020; Hu et al., 2016; Mashburn et al., 2008; NICHD ECCRN, 2002; Phillips et al., 2000; Pianta et al., 2005; Slot et al., 2015).

In this study, we extend the literature on the associations between structural and process features in three principal ways. First, whereas most studies have focused on a limited range of settings in one district or region and thus often observe limited variability in quality features, we examine quality across the full landscape of group-based education and care settings in one state. Second, whereas most researchers rely on broad, global measures of process quality, which may obscure true associations between structures and processes, we consider more specific process features reflecting the moment-to-moment experiences of children and adults in early education and care settings. Third, whereas traditional analyses tend to examine numerous bivariate associations between individual structures and processes, introducing the risk of overinterpreting associations that occur by chance, we use permutation testing to understand which structures are consistently associated with the collection of process features under consideration.

After reviewing the literature on the link between structural and process features, we document quality features in community-based child care (CCC), family child care (FCC), Head Start (HS), and public school prekindergarten (PSP) programs across Massachusetts. Although quality improvement efforts tend to operate at the systems level by regulating quality across this range of group-based settings, there is little empirical work simultaneously documenting quality features in all of these settings. Next, we examine the links between structural and process quality features in this unique sample. In doing so, this work advances our understanding of quality in early education and care, providing more rigorous and precise information on which to draw in policy and practice discussions tied to quality improvement.

### *Associations Between Structural and Process Quality Features*

Structural quality includes a large and diverse set of features at the systems, center, and classroom levels, all believed to influence what children and adults do every day in early education and care settings (Cryer et al., 1999; Slot, 2018). In this article, we focus on six classroom-level structures, reflecting staff qualifications (years of experience and education), group size, child-to-adult ratio, and curricular materials (use of a formal or social-emotional curriculum), that are commonly regulated and considered directly proximal to process features, which, as defined above, include the everyday activities and interactions in settings thought to drive children's development. Classroom structures are theorized to influence high-quality processes through two primary mechanisms: (1) educator knowledge and (2) educator capacity (Cryer et al., 1999; Phillipsen et al., 1997). Greater staff qualifications, including higher education and more years of experience, are thought to afford educators increased knowledge of what works best to support children's development, in turn increasing the quality of processes in their settings (Burchinal et al., 2002; Early et al., 2007; Lin & Magnuson, 2018). Although the majority of early education settings have multiple adults present, most studies, including ours, focus on the qualifications of a single primary educator, with the understanding that they frequently guide day-to-day operations in settings. As with staff qualifications, having a curriculum is hypothesized to provide educators with an understanding of what to do with children, shaping the nature and content of the activities they plan and how they engage and interact with children during those activities (Jenkins et al., 2019). Finally, group size and child-to-adult ratio are thought to influence educators' capacity to engage in rigorous, dynamic activities and emotionally supportive and responsive interactions with children, as those in settings with relatively few children may have more time and headspace to plan and execute these types of processes than those in settings with more children (NICHD ECCRN, 2002).

A large body of empirical research has explored whether these structural features are indeed linked to a wide range of process features (e.g., Burchinal et al., 2002; Cryer et al., 1999; Early et al., 2006; Early et al., 2007; Hanno et al., 2020; Hu et al., 2016; Lin & Magnuson, 2018; Mashburn et al., 2008; NICHD ECCRN, 2002; Phillips et al., 2000; Phillipsen et al., 1997; Pianta et al., 2005; Slot et al., 2015; Slot et al., 2018). To address this question, most studies have tended to apply bivariate (e.g., correlations) and regression-based approaches to document associations between a set of structures and a set of processes. For example, in a multistate study of 238 PSP classrooms, Pianta and colleagues (2005) examined the associations between nine teacher and program features and seven process quality features, including the nature of teacher–child interactions and time spent in different activity formats (e.g., whole group, free-choice centers), for a total of more than 60 structure-process comparisons. For the most part, studies applying these types of methods have been inconclusive, with only some finding statistically significant associations between educator education level, educator experience, child-to-adult ratios, or group size, on one hand, and process quality features, on the other (see Appendix Table A1 for a summary of findings from the cited literature). Even within studies, associations between a given structure and the range of process features considered are often inconsistent. Pianta and colleagues, for example, found that educator experience was related with two of the seven process features they considered.

There are fewer studies examining whether curricular materials are associated with process features, although a number of studies evaluate the impact of a specific curriculum on educator practices, which, like process features, reflect how educators

engage and interact with children (e.g., Barnett et al., 2008; Clements & Sarama, 2008; Domitrovich et al., 2009). For example, Domitrovich and colleagues (2009) found that HS educators randomly assigned to receive a curriculum and associated training supports focused on language, literacy, and social-emotional development spoke more with children and engaged in deeper conversations with them than those who did not receive the supports. Importantly, educators in these types of evaluation studies typically receive extensive training and ongoing mentorship (e.g., in the Domitrovich et al., 2009, study, the educators had 4 days of workshops focused on the curriculum and regular support from a mentor teacher). As such, these studies provide little insight into how curricula might influence educator practices or process features in more typical conditions. Moreover, they tend to evaluate curricula focused on learning in a specific domain, like math or literacy, offering little evidence on the role of whole-child curricula focused on supporting children's development across a number of domains. To address this gap, Jenkins and colleagues (2019) examined the variation in processes between classrooms where teachers were or were not using a comprehensive curriculum. They found that 8 of 13 bivariate comparisons between process features reflecting the quality of teacher-child interactions in classrooms with and without curricula were statistically significant, although few were robust to more rigorous regression-based approaches. Taken together, there is little consistent evidence that process features are more prevalent in classrooms where curricula are being used than in those where they are absent.

In general, the number of associations typically tested in studies linking structures and processes makes it challenging to know whether significant associations are the product of random statistical chance or whether children in settings defined by certain structures indeed have meaningfully different day-to-day experiences than those in other settings. Although seldom used in this literature, validation tests for multiple hypotheses based on resampling techniques—like permutation testing—consider structures' associations with multiple processes simultaneously and offer an opportunity to learn whether a given structural feature is generally associated with a set of process quality features above and beyond what is expected by random chance (Sherman & Funder, 2009). Such approaches advance our knowledge base by moving beyond a singular focus on individual associations between specific structures and specific processes to rigorously evaluate the full set of associations holistically.

### *Moving From Molar to Molecular Features of Process Quality*

Beyond the preponderance of bivariate comparisons, researchers have largely relied on global, or molar, process measures that may obscure the associations between structural features and the elemental, or molecular, processes constituting children's moment-to-moment experiences in early education and care settings.[1] Specifically, the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) and the Early Childhood Environment Rating Scales (ECERS; Harms & Clifford, 1980) are used in the majority of relevant studies, and these are both global observational measures in the sense that their scores reflect observers' summative assessment across multiple theoretically motivated, abstract properties of processes (Pianta et al., 2020). For example, the CLASS' three domains represent composites of multiple processes. More concretely, the instructional support domain includes how well educators engage in rigorous instruction and model language, among other practices (Hamre, 2014). This means two settings may have identical instructional support scores but one might have high levels of rigorous instruction and low language supports, whereas the other might lack rigorous instruction and have rich language supports. Blending these more fine-grained processes is problematic if they each have unique associations with structures. For example, it may be that group size is related to educators' capacity to engage in rigorous instruction but not to how they model language. In this example, when the two processes are combined, the association between group size and instructional support would likely be muted.

Novel approaches to operationalizing process quality in early education and care settings offer insight into the everyday by capturing more precise, molecular features of process quality reflecting discrete behaviors. Specifically, the Teacher Observation in Preschool (TOP; Bilbrey et al., 2007) protocol takes repeated, seconds-long snapshots—or sweeps—of individual educators' behaviors to precisely characterize how they spend their time with children. Over the course of 3 to 5 seconds, observers watch a specific educator and note what they are doing in that moment along six dimensions: language use (e.g., talking, listening), schedule (e.g., whole group, centers), task (e.g., instruction, behavior approving/disapproving), level of instruction, focus (e.g., math, English language arts [ELA]), and tone. For example, while observing an educator for the dimension of language use, observers note whether they are talking, listening to another person, or neither. If the educator is talking or listening, the observer also records to whom (e.g., a child or another adult). At the end of the observation, these snapshots are averaged to yield the proportion of time adults in a classroom are engaged in language-rich interactions with children.

In partnership with a PSP program, the creators of the TOP identified what they believed to be the most salient processes for children's positive developmental outcomes recorded by the TOP[2] (Farran et al., 2017). They argued that high-quality settings were marked by (1) high levels of instruction, (2) more positive emotional climates (i.e., positive tone, little behavior disapproving and much behavior approving), and (3) a high frequency of educators listening to children. Beyond these processes, the TOP captures many additional elemental processes that are likely important to children's learning and development. For instance, some research suggests that how

frequently educators talk with and model language for children may relate to children's own language development (Justice et al., 2008; Wasik et al., 2006). How educators spend their time with children may also matter for a broader range of child skills (Camilli et al., 2010; Fuhs et al., 2013). Specifically, the TOP captures the schedule of settings, meaning whether educators are engaged in whole-group activities, in center-based activities, or in transition. It also records the tasks educators are engaged in, such as instructional activities (i.e., any learning activity during which an educator is engaging with a child) or supporting the personal care needs of children (e.g., tying shoes, blowing noses). Finally, observers note the content area focus of educators' time, including ELA, math, and science. Importantly, while the amount of time spent in any given activity type or content area may be a regulated structure (e.g., a mandated 20-minute math block or an hour in free-choice centers), we consider the amount of time *actually* spent in that content area as a dynamic process reflecting the activities educators and children engage in (i.e., even with a mandated block, time actually spent on the relevant content area is likely to vary).

To date, little research has examined the associations between structures and these more molecular aspects of process quality. As hypothesized in prior studies, we might expect educators with more years of education or experience, in settings with fewer students or lower child-to-adult ratios, or with curricula to enact more desirable processes (e.g., higher instructional levels, more positive tone) and with greater frequency (e.g., more time on instruction and focused on academic content, more talking and listening, and more behavior approving) than educators with less experience or education, in contexts with more students or higher ratios, or without curricula. We might also expect less desirable process quality features (e.g., more time in personal care activities, transitions, or behavior disapproving) to occur with less frequency in settings led by educators with relatively more education or experience, with fewer students or lower ratios, or with curricula. There are other processes, such as time spent in whole groups or centers, whose hypothesized relations to structural features are less clear. It is also possible that there exist differential associations between specific structures and specific processes. For example, curricular materials might be most closely related to time spent in instruction and on academic content given that such materials often include explicit guidance about time usage. However, in the absence of extensive research on molecular process features, these kinds of differential associations are poorly understood.

### *The Current Study*

Building on the robust literature on quality in early education and care, our first goal is to describe the structural and process features of settings across the contemporary landscape of early education and care. To address this aim, we collected detailed information about quality features in CCC, HS, FCC, and PSP programs across the state of Massachusetts. Our second objective is to examine the associations between structural and process quality features. In addressing these two aims, we extend the literature on quality in early education and care in three principal ways. First, we consider the full range of group-based settings where young children spend their time. This allows us to observe greater variation in structural features than is typical in studies that focus on only one type of early education and care setting. Second, we consider more molecular features of process quality than are captured by commonly used global measures, shedding light on the precise moment-to-moment processes that children experience in early education and care settings. Third, in addition to applying traditional bivariate approaches, we use a methodological approach rarely applied in this area of research to generate insight into which structural features are most consistently associated with process quality. In doing so, we identify structures that are associated with specific facets of children's early learning experiences, as well as with meaningful differences in the full range of process features.

## Method

### *Sample*

Data for the current study come from the first year of the Early Learning Study at Harvard (ELS@H; Jones et al., 2020), which was designed to be representative of 3- and 4-year-old children living in Massachusetts. Using three primary recruitment methods (i.e., a household survey, network sampling, and random sampling of licensed early education and care settings), the sample in the first year of the study included 3,222 children (see Jones et al., 2020, for additional details on recruitment). The early education and care providers of these children, the focus of the present study, were also recruited to participate in the study.

The present study's analytic sample comprised all 672 classrooms in the 451 group-based programs that were observed during the first year of the study. This sample includes classrooms in all four types of group-based programs (i.e., CCC, HS, FCC, and PSP) located in 247 distinct zip codes, representing almost half the zip codes in the state. The majority of educators leading these classrooms were female (98.49%), spoke English as a primary language (89.94%), and were White (78.07%). Relatively fewer educators in the sample were Latinx (13.23%), Asian (6.43%), or Black (3.78%). On average, the educators in the sample were 45.11 years old ($SD = 11.93$). There was wide variation in the household incomes of the educators. (See Table 1 and Appendix Table A2 for additional descriptive information about the educators in the analytic sample.) Given the study's sampling approach, demographic information for most of the children in these settings was not available, although the programs were in communities reflecting the socio-demographic diversity of Massachusetts (see Appendix Table A3).

TABLE 1
*Descriptive Characteristics of Educators (N = 672)*

| Characteristic | *n* | *M*/% | *SD* | Min. | Max. |
|---|---|---|---|---|---|
| Age, years | 522 | 45.11 | 11.93 | 19.00 | 71.00 |
| Sex ("1" = female) | 529 | 98.49% | | | |
| Race/ethnicity | | | | | |
| Asian | 529 | 6.43% | | | |
| Black | 529 | 3.78% | | | |
| Latinx | 529 | 13.23% | | | |
| White | 529 | 78.07% | | | |
| Primary language English | 527 | 89.94% | | | |
| Household income ($) | | | | | |
| <30,000 | 495 | 10.71% | | | |
| 30,001–75,000 | 495 | 40.20% | | | |
| 75,001–125,000 | 495 | 30.51% | | | |
| 125,001–200,000 | 495 | 14.34% | | | |
| >201,000 | 495 | 4.24% | | | |
| Provider type | | | | | |
| Community-based child care | 672 | 43.45% | | | |
| Family child care | 672 | 19.64% | | | |
| Head Start | 672 | 19.20% | | | |
| Public school prekindergarten | 672 | 17.71% | | | |

### *Procedure*

Data were collected through educator surveys and in-person observations. Parents of ELS@H participants were asked to provide information on their child's educator.[3] These educators were then asked to complete an online survey between April and August 2018 with a wide variety of items, including about their demographic characteristics, qualifications (e.g., education and experience levels), and practices (e.g., use of a curricula). In a few cases ($n$ = 26), the educators completed the survey on paper. During the same period (April to August 2018), trained observers conducted in-person visits in early education and care settings. The visits typically lasted 4 hours, beginning in the morning and continuing until nap time in the afternoon.

### *Measures*

*Features of Structural Quality.* On the survey, the educators reported their years of experience and education level. Specifically, they were asked to report the number of years they had been taking care of or teaching children (excluding their own). They were also asked to report the highest grade or year of schooling completed from a list of options (i.e., 12th grade or below, high school diploma/GED, vocational/technical program, some college, associate's degree [2-year degree], bachelor's degree [4-year degree], some graduate or professional school, master's degree, or doctoral degree). For the analyses, education level was converted into a continuous variable (e.g., those with a high school diploma were treated as having 12 years of education).

The educators also reported on the survey whether they used curricula in their programs. First, they were asked whether a published curriculum was a source of learning activities in their program. From this item, a binary indicator was created to distinguish between programs using a formal curriculum and those that were not. Second, the educators were asked whether they or their program used a curriculum to "help children learn about their own emotions and other children's emotions and about managing their own behavior." Educators who responded affirmatively to this item were regarded as using a social-emotional curriculum.

Finally, group size and child-to-adult ratio were obtained during the in-person observations. Group size was the total number of children who were observed in the classroom, and child-to-adult ratio was the number of children observed per educator observed in the classroom. Child-to-adult ratio was calculated by dividing the number of children in the setting by the number of educators.

*Features of Process Quality.* Information on 14 features of process quality was captured using the TOP (Bilbrey et al., 2007) protocol. During the in-person visits, trained observers conducted repeated sweeps, each lasting typically 3–5 seconds, for the duration of the observation (approximately 4 hours across all settings). As part of each sweep, the observers noted the specific behaviors of an educator in that moment. In case there were multiple educators in a setting, one educator was observed for several seconds, then the next educator was observed for the next few seconds, and so on. A subset of observed behaviors served as our 14 process quality measures.

These observed behaviors included those related to the educators' use of language—that is, whether (1) listening to children or (2) talking with children; the format of the current activity or schedule—that is, whether in (3) whole group, (4) transitions, or (5) centers; the type of task they were completing—that is, whether (6) approving or (7) disapproving of children's behavior, (8) engaging in instruction, or (9) supporting the personal care needs of children; and the domain focus of the current activity or task—that is, whether focused on (10) ELA, (11) math, or (12) science. The observers also noted the (13) average tone of the adult's voice in that moment, using a 5-point scale ranging from 1 = *extremely negative* to 5 = *vibrant*. If educators were engaged in instruction, observers also rated the (14) level of instruction, using a 4-point scale ranging from 1 = *low-level instruction* to 4 = *high inferential learning*. Across the 672 classrooms in the analytic sample, these features were observed for 1,413 unique adults for an average of 15.28 sweeps each ($SD$ = 3.82).

After receiving training on the TOP codebook and observation protocol, all observers were required to demonstrate high levels of agreement with master observers during in-person practice observations. Agreement between the trainees' codes and those of the master observers was calculated to establish reliability. Across all TOP items, average agreement was 89%.

Classroom-level scores for each feature were constructed by averaging the counts or ratings from all educator sweeps in the classroom. In the case of features that were marked during sweeps as being present or not (e.g., whether or not the adult was talking to children), the average of counts across all sweeps represents the proportion of sweeps during which the behavior was present. In the case of features rated using a scale (i.e., tone and instructional level), the average of sweep-level ratings represents the average level across all sweeps.

*Covariates.* Information on educator demographics (i.e., race/ethnicity, primary language, and household income) came from the online survey. Information on provider type (i.e., whether a program was a CCC, FCC, HS, or PSP program) came from administrative records.

### *Analytic Plan*

First, to understand the structural and process features across the landscape of early education and care in Massachusetts, we documented the levels of structural and process features in the sample. Second, to assess the associations between the structural and process features, we initially applied traditional bivariate approaches consistent with those employed in previous studies. Specifically, we examined the correlations between each structural and process feature. We then used multilevel models with program-level random intercepts to predict each of the 14 features of process quality separately as a function of educator years of experience, educator years of education, group size, child-to-adult ratio, an indicator for whether a formal curriculum was used, and an indicator for whether a social-emotional learning curriculum was used (a total of 14 models). Because structural features were likely to differ systematically across provider types (see Appendix Table A4), all models included provider type controls. Additional covariates included the set of educator demographics outlined above. We adopted program-level random intercepts to account for the nesting of classrooms within settings, as more than one third of the programs (162 of the 451 programs; 35.92%) had multiple classrooms observed.

Recognizing that the number of individual comparisons made using traditional approaches increases the risk of interpreting spurious associations, we then used permutation testing to determine whether the identified associations were more than what was expected by chance. In doing so, we also identified which structures were consistently associated with the range of process features considered. Following the procedures outlined by Sherman and Funder (2009), the six structural characteristics were randomly redistributed across the classrooms in the sample, and then multilevel models predicting *z*-scored process quality features from the structural features with covariates were run within the pseudosamples. Redistributing the structural features across classrooms generates a null association between the predictors (structures) and outcomes (processes) while maintaining the observed associations between covariates and outcomes. This randomization and reestimation procedure was repeated 1,000 times. Then, we calculated the average absolute associations between each structural quality feature and the 14 process quality features across these 1,000 permutations. For each structure, the distribution of the average absolute associations from the permutations—which approximates the sampling distribution of the average absolute association expected by chance—was then charted and compared with the observed average absolute association in the analytic sample. The proportion of absolute associations from the permutations greater than the absolute observed association was then calculated for each structural feature. Smaller proportions indicated that the absolute association between that structure and the 14 processes in the observed sample was greater than expected by chance, whereas larger proportions indicated that the observed absolute association could have been expected by chance.

*Missing Data.* Given our focus on settings with observation data, there was no missingness on process quality features, with the exception of instructional level. In 7 classrooms (1.04%), the educators did not engage in any instruction during the entire observational period, and therefore the level of instruction in those settings was never assessed. These missing instructional quality observations were not imputed for analysis. There was relatively more missingness on the structural features (ranging from 0.00% to 23.96%) primarily because not every classroom had an educator complete the survey. Of the 672 classrooms in the analytic sample, 537 (79.91%) had survey data. As educators were not required to complete all items of the survey, there was also some missingness on particular items among those educators who did take the survey (see Tables 1 and 2 for

TABLE 2
*Descriptive Statistics of Features of Structural and Process Quality*

| Feature | *n* | *M*/% | *SD* | Min. | Max. |
|---|---|---|---|---|---|
| *Features of structural quality* | | | | | |
| Years of experience | 524 | 18.51 | 10.14 | 1.00 | 47.00 |
| Education level | 527 | 15.76 | 1.88 | 10.00 | 20.00 |
| Group size | 672 | 10.82 | 5.00 | 1.00 | 24.00 |
| Child-to-adult ratio | 672 | 5.28 | 2.37 | 0.50 | 16.00 |
| Uses a formal curriculum | 537 | 45.07% | | | |
| Uses a social-emotional curriculum | 511 | 80.04% | | | |
| *Features of process quality* | | | | | |
| Language use | | | | | |
| Proportion of sweeps listening to children | 672 | 6.27 | 6.54 | 0.00 | 40.00 |
| Proportion of sweeps talking to children | 672 | 57.21 | 15.58 | 18.52 | 100.00 |
| Schedule | | | | | |
| Proportion of sweeps in whole group | 672 | 31.45 | 17.86 | 0.00 | 100.00 |
| Proportion of sweeps in transition | 672 | 14.87 | 9.98 | 0.00 | 48.48 |
| Proportion of sweeps in centers | 672 | 27.40 | 20.16 | 0.00 | 100.00 |
| Task | | | | | |
| Proportion of sweeps behavior approving | 672 | 2.91 | 3.81 | 0.00 | 25.00 |
| Proportion of sweeps behavior disapproving | 672 | 7.82 | 7.81 | 0.00 | 56.25 |
| Proportion of sweeps in instruction | 672 | 31.47 | 16.11 | 0.00 | 100.00 |
| Proportion of sweeps in personal care activities | 672 | 12.48 | 10.19 | 0.00 | 47.50 |
| Instructional level | | | | | |
| Average instructional level | 665 | 1.65 | 0.33 | 1.00 | 2.41 |
| Focus | | | | | |
| Proportion of sweeps in English language arts | 672 | 8.84 | 8.07 | 0.00 | 50.00 |
| Proportion of sweeps in math | 672 | 3.63 | 5.37 | 0.00 | 45.00 |
| Proportion of sweeps in science | 672 | 5.32 | 7.13 | 0.00 | 80.00 |
| Tone | | | | | |
| Average tone | 672 | 3.41 | 0.32 | 2.38 | 4.62 |

details). For the multilevel regression analyses, missing covariates and structural features were imputed using multiple imputation by chained equations in Stata 17 (StataCorp, 2021), a process that yielded 20 complete data sets. The first complete data set from this procedure was used in the permutation testing given its computational intensiveness. Complete-case analysis indicated that the findings from the multilevel modeling and permutation testing were not sensitive to our missing data approach (see Appendix Table A5 and Appendix Figure A1).

## Results

### *Features of Structural and Process Quality*

*Structural Quality.* The top panel of Table 2 presents descriptive information on features of structural quality. Educators across Massachusetts had an average of nearly two decades of experience working with young children ($M = 18.51$ years, $SD = 10.14$). They also had an average of 15.76 years of education ($SD = 1.88$), equivalent to slightly less than a bachelor's degree (16 years). The classrooms in the sample had an average of 10.82 children ($SD = 5.00$) and an average child-to-adult ratio of 5.28 ($SD = 2.37$), indicating that many had multiple adults present. The histograms in Figure 1 illustrate variability across these four continuous structural features (i.e., educator experience, educator education, group size, and child-to-adult ratio). Finally, almost half the educators (45.07%) reported using a formal published curriculum to guide learning activities, and far more educators reported using a social-emotional curriculum (80.04%).

*Process Quality.* Table 2 also presents descriptive information on the 14 process quality features we considered in this study. In terms of the language environment, educators tended to speak to children in the majority of sweeps ($M = 57.21$, $SD = 15.58$). Educators tended to listen to children far less than they spoke to children ($M = 6.27$, $SD = 6.54$).

Features captured by the schedule, task, and focus variables suggest that the educators engaged in a diverse range of activities throughout the day. Educators spent approximately

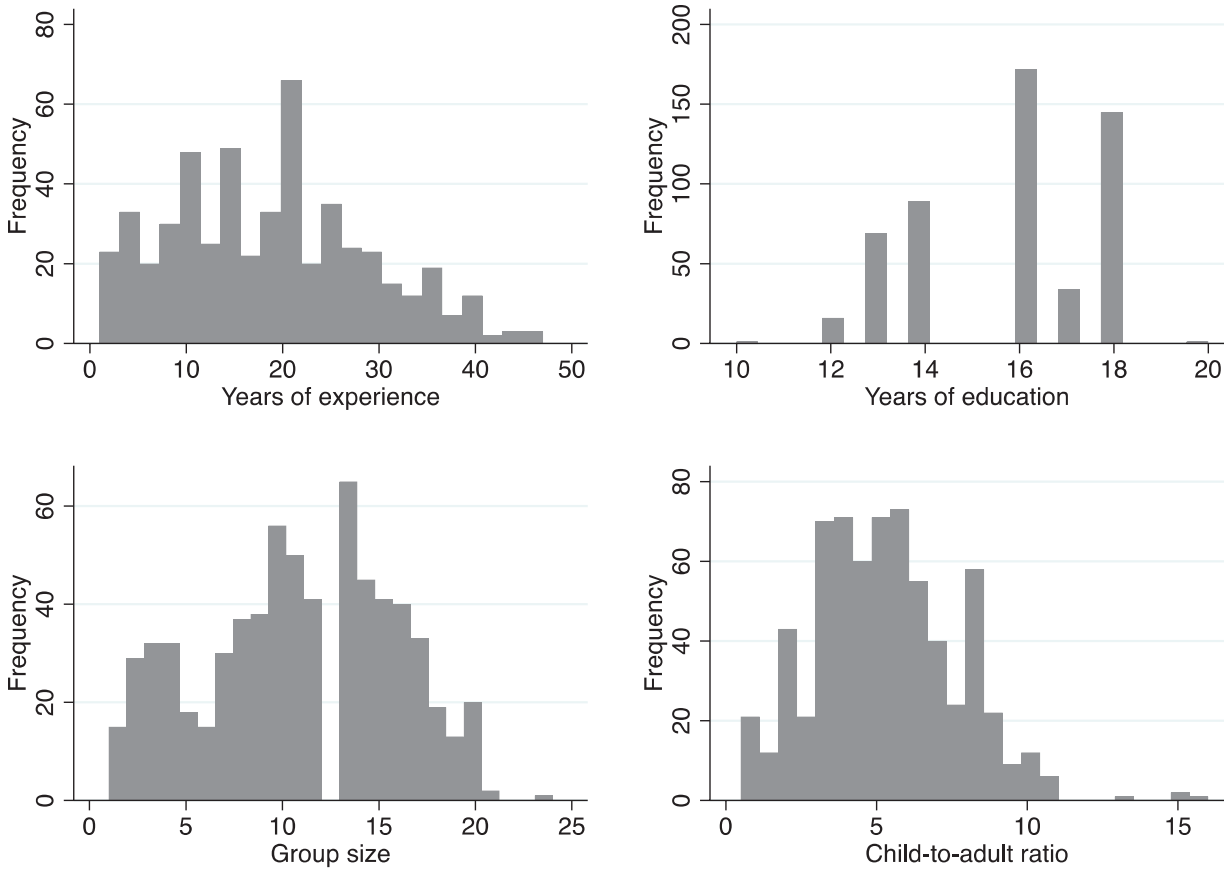


FIGURE 1. *Distribution of continuous structural features.*

30% of sweeps in whole groups and centers and an average of 14.87% of sweeps in transition ($SD = 9.98$). In terms of the types of tasks the educators engaged in, instructional tasks took place in approximately 30% of the sweeps ($M = 31.47$, $SD = 16.11$), although a considerable proportion of sweeps were also spent attending to the personal care needs of children ($M = 12.48$, $SD = 10.19$). Educators spent less time approving or disapproving of children's behaviors than they did in instruction or on personal care. Relatively more sweeps were spent disapproving of children's behaviors ($M = 7.82$, $SD = 7.81$) than approving of children's behaviors ($M = 2.91$, $SD = 3.81$). In terms of focus, educators spent little time focused on traditional academic domains (i.e., ELA, math, and science). They spent approximately one tenth of sweeps focused on ELA ($M = 8.84$, $SD = 8.07$). Educators tended to spend less time focused on math ($M = 3.63$, $SD = 5.37$) or science ($M = 5.32$, $SD = 7.13$) than they did on ELA.

The average instructional level across the sample was low ($M = 1.65$, $SD = 0.33$), indicating that instruction was most often either low-level instruction (e.g., cutting with scissors, singing songs) or focused on basic skills (e.g., learning letters and numbers). Average tone in the sample was 3.41 ($SD = 0.32$), suggesting that the adults' tone tended to be somewhere between flat and pleasant.

### *Linking Structural and Process Quality Features*

Table 3 presents bivariate correlations between structural and process quality features, and Table 4 presents multilevel models predicting each process feature from the structural quality indicators, accounting for covariates. In general, results from these two approaches suggest that group size and child-to-adult ratio were most consistently and robustly associated with the array of process features. Specifically, group size was a significant predictor of 6 of the 14 process features, accounting for covariates, and child-to-adult ratio was a significant predictor of 5. As compared with group size or child-to-adult ratio, the remaining structural features were less consistently linked to process quality features. We noted that even in the presence of some statistically significant associations between structural and process features, structural features appeared to account for little variation in process features across classrooms (see Appendix Table A6).

Group size was associated with process features capturing language use, schedule, task, focus, and tone. In large groups, educators tended to talk less to children. Each additional child in the classroom was associated with 1.68 percentage points ($SE = 0.19$, $p < .001$) fewer sweeps with adults talking with children, accounting for the other structural features, educator covariates, and provider type. Educators in larger groups also

TABLE 3
*Bivariate Correlations Between Features of Structural and Process Quality*

| | Language use | | Schedule | | | Task | | | | Instructional level | Focus | | | Tone |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Feature | Sweeps listening to children | Sweeps talking to children | Sweeps in whole group | Sweeps in transition | Sweeps in centers | Sweeps behavior approving | Sweeps behavior disapproving | Sweeps in instruction | Sweeps in personal care activities | Average instructional level | Sweeps in English language arts | Sweeps in math | Sweeps in science | Average tone |
| Years of experience | −0.04 | 0.13** | −0.02 | −0.06 | 0.05 | 0.02 | −0.06 | 0.06 | 0.00 | 0.05 | 0.05 | −0.04 | 0.16*** | 0.04 |
| Years of education | −0.07 | −0.02 | 0.04 | 0.02 | −0.03 | −0.09* | −0.15*** | 0.08 | −0.09* | −0.03 | 0.04 | 0.00 | −0.03 | 0.09* |
| Group size | −0.02 | −0.34*** | −0.08* | 0.17*** | −0.03 | −0.10** | 0.08* | −0.23*** | −0.05 | 0.02 | −0.07 | −0.16*** | −0.20*** | −0.10* |
| Child-to-adult ratio | 0.00 | −0.02 | −0.01 | 0.11** | −0.03 | −0.06 | 0.14*** | −0.03 | −0.06 | 0.09* | 0.11** | −0.08* | −0.12** | −0.01 |
| Uses a formal curriculum | −0.02 | −0.08 | 0.09* | 0.02 | −0.12** | 0.06 | −0.05 | 0.00 | −0.12** | −0.03 | 0.03 | −0.02 | −0.00 | 0.00 |
| Uses a social-emotional curriculum | 0.01 | −0.03 | 0.11* | 0.09 | −0.10* | 0.03 | 0.05 | −0.05 | −0.05 | 0.01 | −0.05 | 0.02 | −0.03 | −0.04 |

*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 4
*Multilevel Models Predicting Process Quality From Structural Quality*

| | Language use | | Schedule | | | Task | | | | Instructional level | Focus | | | Tone |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Feature | (1) Sweeps listening to children | (2) Sweeps talking to children | (3) Sweeps in whole group | (4) Sweeps in transition | (5) Sweeps in centers | (6) Sweeps behavior approving | (7) Sweeps behavior disapproving | (8) Sweeps in instruction | (9) Sweeps in personal care activities | (10) Average instructional level | (11) Sweeps in English language arts | (12) Sweeps in math | (13) Sweeps in science | (14) Average tone |
| Years of experience | −0.01 (0.03) | 0.10 (0.06) | −0.05 (0.07) | 0.00 (0.04) | 0.07 (0.09) | −0.01 (0.02) | −0.00 (0.03) | −0.02 (0.06) | 0.02 (0.04) | 0.00 (0.00) | 0.02 (0.03) | −0.03 (0.03) | 0.05 (0.03) | −0.00 (0.00) |
| Years of education | −0.16 (0.17) | 0.41 (0.40) | 0.18 (0.46) | 0.14 (0.26) | 0.49 (0.55) | −0.14 (0.11) | −0.18 (0.19) | 0.13 (0.40) | −0.09 (0.25) | −0.01 (0.01) | −0.04 (0.20) | −0.02 (0.15) | −0.05 (0.20) | −0.00 (0.01) |
| Group size | 0.01 (0.09) | −1.68*** (0.19) | −0.24 (0.23) | 0.28* (0.13) | −0.14 (0.26) | −0.01 (0.05) | −0.02 (0.10) | −1.15*** (0.20) | 0.06 (0.13) | 0.00 (0.00) | −0.37*** (0.10) | −0.14* (0.07) | −0.14 (0.09) | −0.01* (0.00) |
| Child-to-adult ratio | 0.06 (0.15) | 2.94*** (0.33) | 0.92* (0.40) | −0.10 (0.23) | −0.31 (0.47) | 0.06 (0.09) | 0.48** (0.18) | 1.87*** (0.35) | −0.41 (0.23) | 0.01 (0.01) | 1.03*** (0.18) | 0.22 (0.13) | 0.03 (0.17) | 0.01 (0.01) |
| Uses a formal curriculum | −0.13 (0.56) | −1.46 (1.20) | 1.24 (1.43) | −0.10 (0.89) | −2.22 (1.70) | 0.29 (0.34) | −0.07 (0.66) | 0.01 (1.26) | −1.29 (0.86) | −0.02 (0.03) | 0.32 (0.65) | −0.06 (0.49) | 0.13 (0.60) | −0.01 (0.02) |
| Uses a social-emotional curriculum | 0.36 (0.65) | 0.40 (1.49) | 2.63 (1.74) | 1.51 (1.09) | −3.09 (2.08) | 0.39 (0.41) | 0.28 (0.83) | −0.40 (1.47) | 0.09 (1.02) | 0.03 (0.03) | −0.19 (0.86) | 0.16 (0.61) | 0.04 (0.80) | −0.00 (0.03) |
| *N* | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 672 | 665 | 672 | 672 | 672 | 672 |

*Note.* All models include a program-level random intercept. All models include indicator variables for whether the program was a public school prekindergarten, Head Start, or family child care, making community-based child care the reference group. Additional covariates in all models include educator household income and indicators for whether they were White or spoke English as a primary language. Standard errors are in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$.

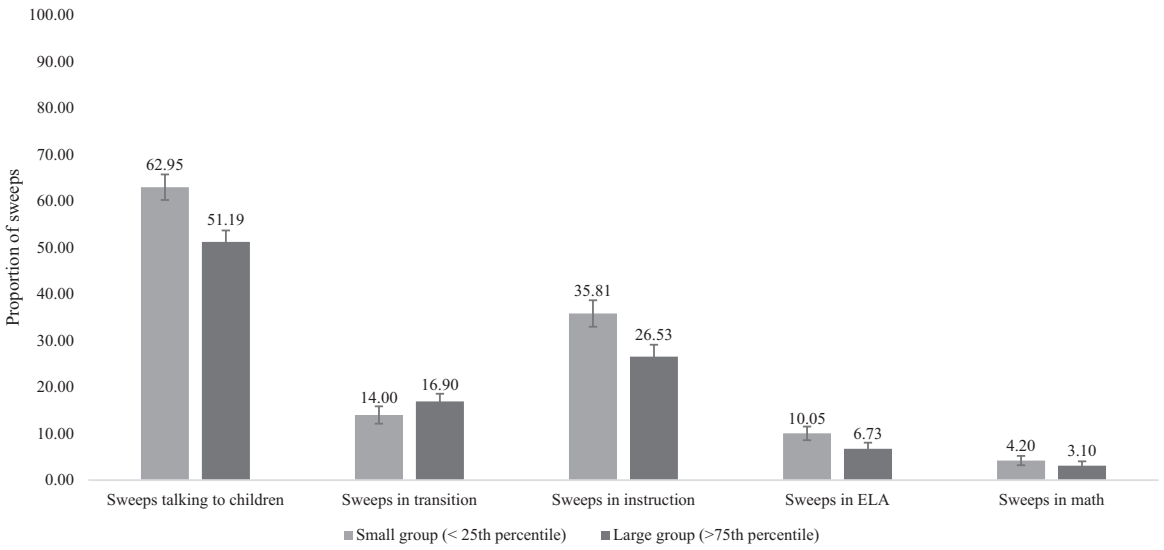


FIGURE 2. *Predicted process quality features by classroom group size.*
*Note.* Estimates come from two separate multilevel regression models: one including an indicator for whether or not the classroom had a group size the 25th percentile or below (i.e., group size of 8 or fewer children) and the other including an indicator for whether or not the classroom had a group size the 75th percentile or above (i.e., group size of 15 or more children). All models account for other structural features (i.e., years of experience, years of education, child-to-adult ratio, use of a formal curriculum, use of a social-emotional curriculum), educator demographics, and provider type. The bars represent ±1.96 standard error of the estimates. The process features included had at least one statistically significant association at $p < .05$ in Table 4, with the exclusion of educator tone, which is on a different scale. ELA = English language arts.

tended to spend more time in transition and less time engaged in instructional tasks than those in smaller groups. More concretely, each additional child was associated with 0.28 percentage point ($SE = 0.13$, $p < .05$) more sweeps spent in transition and 1.15 percentage points ($SE = 0.19$, $p < .001$) fewer sweeps spent in instructional activities. Group size additionally appeared to be inversely associated with the proportion of time educators spent focused on ELA or math, as well as with average educator tone, although the association between group size and educator tone was substantively small ($b = -0.01$, $SE = 0.00$, $p < .05$). Figure 2 illustrates the predicted moment-to-moment experiences in smaller groups (i.e., 8 or fewer children; 25th percentile for group size) as compared with larger groups (i.e., 15 or more children; 75th percentile for group size). For example, educators in classrooms with 8 or fewer children were estimated to speak to children almost two thirds of the time, whereas those in classrooms with 15 or more children were estimated to speak to children about half of the time.

Although child-to-adult ratio was also related to a number of process features, it was not always associated in expected ways. In line with our hypotheses, behavioral disapproval was more common in settings with larger child-to-adult ratios than in settings with smaller ratios. Adjusting for covariates, each one-unit increment in the child-to-adult ratio was associated with a 0.48 percentage point ($SE = 0.18$, $p < .01$) difference in the proportion of sweeps spent disapproving of children's behaviors. However, contrary to our hypotheses, we found that child-to-adult ratio was positively and strongly associated with (1) more talk with children, (2) more time spent focused on ELA, and (3) more time in instruction. In these instances, the unadjusted bivariate associations between child-to-adult ratio and the process features were indistinguishable from 0 or only weakly correlated. When accounting for other structural features and provider type, however, the relations were substantively large and positive. Each one-unit difference in child-to-adult ratio was associated with a 2.94 percentage point difference in sweeps talking to children ($SE = 0.33$, $p < .001$), a 1.03 percentage point difference in sweeps focused on ELA ($SE = 0.18$, $p < .001$), and a 1.87 percentage point difference in sweeps on instruction ($SE = 0.35$, $p < .001$).

The results of the permutation tests illustrated in Figure 3 confirmed that associations of both group size and child-to-teacher ratio with process features were greater than expected by random chance. In both cases, the observed average absolute associations between process features and child-to-adult ratio and group size were far larger than the associations yielded by chance. This is illustrated in Figure 3 by the black vertical lines representing the observed absolute average association positioned far to the right of the distribution of associations expected by chance. Permutation tests confirmed that educator education level, experience, and use of curricula were not systematically related to process quality features (i.e., the observed average absolute associations could have been observed by random chance). In the absence of this exercise, we may have been tempted to evaluate the significant bivariate associations observed between these four structural features and a minority of processes in Table 3. Permutation testing

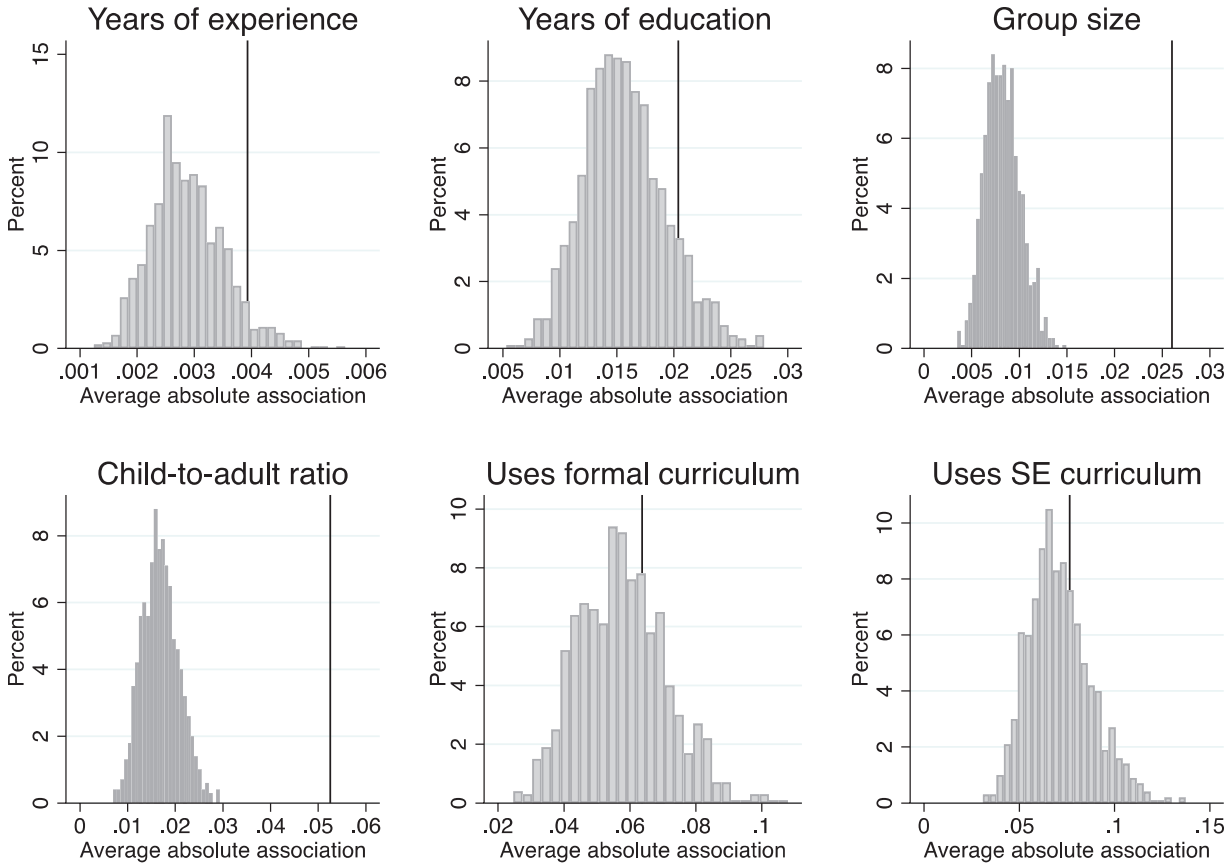


FIGURE 3. *Approximate sampling distribution for the average association between each structural quality feature and process features.*
*Note.* The distribution comes from 1,000 trials with randomized structural features in the first imputed data set. The vertical line indicates the observed absolute average association between the particular structural feature and the 14 features of process quality in the observed data set. All process features were *z* scored prior to the analyses using the permutated data. The models controlled for provider type, educator demographic characteristics, and other structural features.

offers information about the likelihood that associations across the full set of processes were observed by random chance, providing clarity about which bivariate associations to give weight to and, more broadly, which structures were most associated with the range of experiences and activities that constitute process quality.

## Discussion

A large body of research emphasizes the importance of quality for children's learning in early education and care programs (e.g., Hamre, 2014; Hanno et al., 2021; Markowitz et al., 2017; Mashburn et al., 2008), but little work considers the finer-grained, molecular features of process quality across multiple types of early education and care programs, nor has it examined them in relation to oft-regulated structural features. Using unique data from a statewide study of early education and care in Massachusetts, we first documented quality features across the landscape of group-based early education and care in the state. We then examined the associations between structural and process quality features in these settings to inform the conversation on whether certain structures are likely to underlie the day-to-day realities in early education and care programs.

### *Features of Quality in Early Education and Care*

Structural features in our sample mirror broader trends in early education and care. In comparison with educators working with children in the nationally representative Early Childhood Longitudinal Birth Cohort, the educators in our sample had similar average education levels but more years of experience (Bassok et al., 2016). The finding that many educators have worked for several decades with young children is notable given prior work documenting high levels of staff turnover in the early education and care field (Whitebook et al., 2014). It may be that educator turnover in Massachusetts is less of a challenge than in other locations or that educators move between programs within the field. Findings also show that, as hypothesized, early education and care programs tended to have group sizes and child-to-adult ratios lower than are typical in K–12 settings. Last, nearly twice as many educators in our sample reported using a curriculum to guide

social-emotional learning than a formal curriculum to support children's learning across a wider variety of domains. This stands in contrast to the findings of Jenkins and colleagues (2019), which showed that between 60% and 100% of the primarily HS and PSP settings in their sample used a whole-child curriculum. The difference between our findings is likely driven by the presence in our sample of settings in which such curricula are not frequently mandated (e.g., FCC).

The observed process quality features in our sample showed that educators split their time among a variety of activities. First, educators spent far more time talking to children than listening to them. In their analysis of 26 public prekindergarten classrooms, Farran and colleagues (2017) similarly found that educator talk outpaced listening. Second, educators spent the majority of time (58%) engaged in either whole-group or center-based activities, although transitions still accounted for a considerable proportion of time (15%). The combined proportion of time spent in whole groups or centers was nearly identical to that observed by Pianta and colleagues (2005) in their study of PSP programs (59%). Third, in terms of the types of tasks educators completed, educators spent more than twice the proportion of sweeps disapproving of children's behaviors than they did approving of children's behaviors, although the amount of time educators spent doing either was low (<10%). They spent relatively more time in instruction and supporting the personal care needs of children than in behavior approving or disapproving. Finally, less than one fifth of the total time was spent on activities focused on core academic subjects (i.e., ELA, math, or science).

Whereas the aforementioned process features shed light on the moment-to-moment activities of educators (the *what*), the remaining two features—instructional level and tone—provided insight into the qualitative nature of *how* educators operated in their classrooms. The average instructional level in the sample indicated that educators tended to engage mostly in low-level or basic skills instruction, and the average tone indicated that educators tended to use generally pleasant tones. A low average instructional level and a positive average tone were similarly observed in the Farran and colleagues (2017) study. These patterns also align with those found using the CLASS, which show that emotional supports tend to be high across settings but instructional supports tend to be low (Hamre et al., 2007; Mashburn et al., 2008).

### *Linking Structural and Process Quality Features*

We found that group size and child-to-adult ratio were significantly and consistently associated with process quality. This implies that children and adults have different moment-to-moment experiences in settings with a large number of children than in those with a small number of children. It also implies that children have different experiences in contexts where individual adults are responsible for many children than in those where adults are responsible for a few children. Despite the general associations between both group size and child-to-adult ratio with process quality features, we noted that group size appeared more consistently related to process quality than child-to-adult ratio. Although group size and child-to-adult ratio are likely both proxies for the load educators face, group size likely represents the *overall* administrative burden on educators. This burden may be relevant for how educators spend their time if certain activities are challenging to facilitate in a large group regardless of the number of adults present. For example, if children must wash their hands before eating a snack (a personal care task), it is almost certain to take longer with a group of 20 children than with a group of 10 children.

We also observed several unexpected associations suggesting that higher child-to-adult ratios were positively associated with process quality after accounting for the other structural features. Specifically, we found that educators tended to speak more to children, spend more time in instruction, and focus more on ELA in contexts with higher child-to-adult ratios. These results imply, for example, that if we were to compare two classrooms both with 20 students, but one had three educators and the other had two, we would expect the classroom with two educators to focus more on ELA and have more talk with children than the one with three educators. It could be that classrooms with relatively more educators adopt a "divide and conquer" approach, with particular educators devoting more of their time to administrative tasks or children's personal care needs and less time to direct interaction and instruction. More research is needed to understand how educators in settings with multiple adults share and distribute job demands (Sheridan et al., 2014).

Although group size and child-to-adult ratio were associated with process features capturing time allocation, neither was strongly associated with the two qualitative process features (i.e., instructional level and educator tone). Group size was inversely associated with average tone, but the association was substantively small. This suggests that these two structural features might shape *what* activities educators do but may not necessarily influence *how* they engage in those activities. For example, an educator in a large classroom may not be able to talk much with children due to many competing demands but may be able to maintain a pleasant tone while speaking with children.

In contrast to our findings that group size and child-to-adult ratio were significantly related to many process quality features, there was little evidence that educator experience, education level, or curriculum use were consistently related to process quality. As hypothesized in prior work (e.g., Hanno et al., 2020; Slot et al., 2015), the absence of associations between these structural and process quality features could be attributable to several factors. First, it may be that the structural features as measured here obscure the wide variation within like levels. For example, two educators may both have a bachelor's degree, but they could have attended programs that prepared them in unique

ways to work with young children. In line with this hypothesis, some evidence suggests that settings led by educators who do not receive any early childhood–specific credits have lower process quality than those with educators who do have specialized training (Lin & Magnuson, 2018). Relatedly, two settings may both have a social-emotional curriculum, but in one it may be used more extensively than in the other or the content of the two curricula could vary greatly.

Second, structural features could have interactive associations with each other (Slot, 2018). For example, it may be that educators are only able to implement high-quality processes if they have a certain level of education *and* use a curriculum. In line with this hypothesis, the "strongest hope" model for improving quality is thought to combine curricula with ongoing, intensive supports for educators, like coaching (Weiland et al., 2018; Yoshikawa et al., 2013). Third, it could be that educator education and experience, in particular, have nonlinear associations with process features. Some recent research conducted in China on the association between structural and process features found that educator experience was nonlinearly related to global processes (Hu et al., 2017). That is, differences in educator experience were associated with positive differences in process quality but only up to a point. Although not exhaustive, these explanations highlight several avenues for future research on educator education, experience, and curriculum use.

### *Limitations and Future Directions*

Despite the numerous strengths of this study (e.g., a large and diverse contemporary sample, novel measurement of process features), there are several important limitations. First, our design does not allow us to make causal assertions about the associations between structural and process features. For example, although we found that group size was linked to a number of process features, we cannot say that group size *caused* differences in these processes. It may be that other factors correlated with both group size and process features, varying within provider types, like program budgets or workplace stress, may explain the associations.

Second, the structural and process measures employed in this study likely obscure important variation in these features within settings. Regarding the structural measures, staff qualifications come from only one educator per setting and are therefore unlikely to reflect the experiences of all educators in each setting. Similarly, group size and child-to-teacher ratio were observed during classroom visits and may therefore not represent stable conditions over time. Regarding the process measures, it is likely there is variation in what individual educators and children experience within the same setting (e.g., some teachers spoke more or less than others in the same classroom). Research is needed to document this variation and understand its implications for children's learning.

Relatedly, our study does not provide evidence on the relevance of the considered process features for children's positive development, which is an important limitation insofar as quality features primarily matter if they support children's learning. Some previous work offers preliminary evidence on the relevance of these types of molecular features for children's development in a number of domains (Farran et al., 2017; Fuhs et al., 2013; Jones et al., 2020). However, as the measurement of molecular features of process quality is relatively novel, little is known about the precise levels of these features necessary to nurture children's healthy development. Whereas it is commonly assumed that higher scores indicate better quality on global measures than lower scores, the same is unlikely to be true of more molecular measures. For example, although more listening to children is likely to be a desirable process feature, a setting where educators spend 100% of their time listening to children, and no time talking, is unlikely to benefit children's development. Similarly, instruction could be inaccessible and cognitively exhausting if educators were to only engage in high levels of inferential learning. Future research is therefore needed to understand the ideal levels and combinations of various molecular features for promoting children's development across learning domains.

Finally, although this sample includes educators from education and care programs across Massachusetts, the results are not representative at the state level. Participation in our study likely varied systematically by educator characteristics as well as by program features. Relatedly, the early education and care system of Massachusetts may be distinct from those of other states.

## Conclusion

As policymakers continue to pursue quality improvement across the early education and care system, our study underscores the importance of taking into account a wider array of quality features and, in particular, considering more precise, molecular features of process quality that capture the moment-to-moment realities of settings. Applying a synthetic methodological approach with unique data on these features from a statewide study of early education and care, we found that group size and child-to-adult ratio were significantly and consistently associated with a range of process quality features. In contrast, we found little evidence that educator education or experience level or the use of a formal or social-emotional curriculum were associated with these process features. As such, blanket requirements for minimum education requirements or curriculum use are unlikely to yield meaningful improvements in children's day-to-day experiences in early education and care programs. These findings do, however, suggest that group size and child-to-adult ratio are potentially important structural features to consider in developing policies that make a difference for those experiences.

## Appendix

TABLE A1

*Overview of the Structural and Process Quality Empirical Studies Included in the Literature Review*

| | | | Structural features | | | | | |
|---|---|---|---|---|---|---|---|---|
| Authors | Number of settings | Process quality measure | Years of experience | Education level | Group size | Child-to-adult ratio | Uses a formal curriculum | Uses a social-emotional curriculum |
| Burchinal et al. (2002) | 553 | CIS, ECERS/ITERS | 0 | 0,* | | 0,* | | |
| Cryer et al. (1999) | 55 to 288 | CIS, ECERS | 0, * | 0, * | 0, * | 0, * | | |
| Early et al. (2006) | 237 | CLASS, ECERS | | 0, * | | | | |
| Early et al. (2007) | 76 to 887 | ECERS, OCRE | | 0, * | | | | |
| Hanno et al. (2020) | 128 | CLASS | 0 | 0 | | 0 | | |
| Hu et al. (2016) | 162 | CECERS | 0 | 0 | | * | | |
| Hu et al. (2017) | 180 | CLASS | * | * | 0 | * | | |
| Jenkins et al. (2019) | 170–997 | Arnett Caregiver Interaction Scale, ECERS, Teacher Behavior Rating Scale | | | | | 0 | |
| Lin & Magnuson (2018) | 189 | ECERS | | 0, * | | | | |
| Mashburn et al. (2008) | 671 | ECERS, CLASS | | 0, * | 0, * | * | | |
| NICHD ECCRN (2002) | 656–789[a] | OCRE | | * | | * | | |
| Phillips et al. (2000) | 104 | ECERS/ITERS, Assessment Profile for Early Childhood Programs | | 0, * | 0 | * | | |
| Phillipsen et al. (1997) | 228–521 | CIS, ECERS/ITERS, Teacher Involvement Scale | 0, * | 0, * | 0, * | * | | |
| Pianta et al. (2005) | 238 | CLASS, ECERS, Emerging Academics Snapshot | 0, * | 0, * | | 0 | | |
| Slot et al. (2015) | 295 | CLASS | | 0, * | 0 | 0 | | |
| Slot et al. (2018) | 260 | CLASS | 0 | 0, * | 0 | 0 | | |

*Note.* Symbols under structural features reflect general findings in each paper, with blanks indicating the association was not examined, 0 = null association, and * = significant association. Some articles considered additional structural features (e.g., compensation, administrator qualifications) not examined in the current study. CIS = Caregiver Interaction Scale; CECERS = Chinese Early Childhood Environmental Rating Scale; CLASS = Classroom Assessment Scoring System; ECERS = Early Childhood Environmental Rating Scale; ITERS = Infant-Toddler Environmental Rating Scale; OCRE = Observational Record of Caregiving Environment; ECCRN = Early Child Care Research Network.
[a]Analyses were conducted at the child level.

TABLE A2

*Descriptive Characteristics of Educators by Provider Type (N = 672)*

| Characteristic | Overall (*N* = 672) | CCC (*n* = 292) | HS (*n* = 129) | FCC (*n* = 132) | PSP (*n* = 119) |
|---|---|---|---|---|---|
| Age, years | 45.11 (11.93) | 42.00 (12.64) | 44.86 (10.88) | 51.38 (9.91) | 44.55 (10.75) |
| Sex ("1" = female) | 0.98 (0.12) | 0.99 (0.12) | 0.98 (0.14) | 0.98 (0.13) | 0.99 (0.10) |
| Race/ethnicity | | | | | |
| Asian | 0.06 (0.25) | 0.05 (0.21) | 0.14 (0.35) | 0.07 (0.25) | 0.02 (0.14) |
| Black | 0.04 (0.19) | 0.04 (0.20) | 0.07 (0.26) | 0.03 (0.18) | 0.00 (0.00) |
| Latinx | 0.13 (0.34) | 0.08 (0.28) | 0.26 (0.44) | 0.19 (0.39) | 0.04 (0.20) |
| White | 0.78 (0.41) | 0.86 (0.35) | 0.54 (0.50) | 0.72 (0.45) | 0.94 (0.24) |
| Primary language English | 0.90 (0.30) | 0.96 (0.19) | 0.84 (0.37) | 0.77 (0.42) | 0.98 (0.14) |
| Household income ($) | | | | | |
| <30,000 | 0.11 (0.31) | 0.11 (0.31) | 0.21 (0.41) | 0.10 (0.30) | 0.00 (0.00) |
| 30,001–75,000 | 0.40 (0.49) | 0.46 (0.50) | 0.54 (0.50) | 0.35 (0.48) | 0.17 (0.38) |
| 75,001–125,000 | 0.31 (0.46) | 0.27 (0.45) | 0.21 (0.41) | 0.37 (0.49) | 0.41 (0.49) |
| 125,001–200,000 | 0.14 (0.35) | 0.11 (0.32) | 0.04 (0.20) | 0.17 (0.37) | 0.30 (0.46) |
| >201,000 | 0.04 (0.20) | 0.05 (0.22) | 0.00 (0.00) | 0.01 (0.10) | 0.12 (0.32) |

*Note.* Standard deviations are in parentheses. CCC = community-based center care; HS = Head Start; FCC = family child care; PSP = public school prekindergarten.

TABLE A3

*Descriptive Characteristics of Program Communities (N = 451)*

| Characteristic | *N* | *M* | *SD* | Min. | Max. |
|---|---|---|---|---|---|
| Race/ethnicity (% of population) | | | | | |
| Asian | 436 | 5.95 | 7.66 | 0.00 | 56.40 |
| Black | 436 | 9.76 | 14.08 | 0.00 | 92.60 |
| Latinx | 436 | 13.67 | 18.95 | 0.00 | 90.90 |
| White | 436 | 81.21 | 18.93 | 7.70 | 100.00 |
| Poverty rate (%) | 436 | 11.53 | 10.96 | 1.00 | 63.50 |
| Median household income ($) | 436 | 82,275.91 | 35,598.42 | 14,250.00 | 205,074.00 |

*Note.* Data are reported at the Census tract level and come from the American Community Survey.

TABLE A4

*Quality Features by Provider Type*

| Feature | Overall (*N* = 672) | CCC (*n* = 292) | HS (*n* = 129) | FCC (*n* = 132) | PSP (*n* = 119) |
|---|---|---|---|---|---|
| *Structural quality* | | | | | |
| Years of experience | 18.51 (10.14) | 17.37 (10.35) | 17.07 (9.17) | 21.62 (10.56) | 18.84 (9.43) |
| Education level | 15.76 (1.88) | 15.68 (1.76) | 15.35 (1.30) | 14.68 (2.05) | 17.68 (0.69) |
| Group size | 10.82 (5.00) | 12.64 (4.08) | 13.18 (3.36) | 3.50 (1.93) | 11.94 (3.05) |
| Child-to-adult ratio | 5.28 (2.37) | 6.37 (2.20) | 5.66 (1.79) | 2.61 (1.39) | 5.18 (1.81) |
| Uses a formal curriculum | 0.45 (0.50) | 0.36 (0.48) | 0.54 (0.50) | 0.39 (0.49) | 0.63 (0.49) |
| Uses a social-emotional curriculum | 0.80 (0.40) | 0.75 (0.43) | 0.91 (0.28) | 0.75 (0.43) | 0.86 (0.35) |
| *Process quality* | | | | | |
| Language use | | | | | |
| Proportion of sweeps listening to children | 6.27 (6.54) | 6.65 (6.69) | 6.03 (6.40) | 6.51 (7.21) | 5.31 (5.39) |
| Proportion of sweeps talking to children | 57.21 (15.58) | 55.44 (15.51) | 51.68 (11.75) | 66.21 (16.85) | 57.55 (13.78) |
| Schedule | | | | | |
| Proportion of sweeps in whole group | 31.45 (17.86) | 28.51 (16.67) | 28.29 (12.35) | 34.73 (24.75) | 38.47 (13.78) |
| Proportion of sweeps in transition | 14.87 (9.98) | 16.04 (10.18) | 17.67 (9.00) | 12.10 (9.29) | 12.07 (9.96) |
| Proportion of sweeps in centers | 27.40 (20.16) | 28.40 (20.17) | 31.04 (16.38) | 27.68 (24.24) | 20.69 (17.41) |
| Task | | | | | |
| Proportion of sweeps behavior approving | 2.91 (3.81) | 2.52 (3.28) | 2.33 (3.15) | 4.06 (4.80) | 3.20 (4.17) |
| Proportion of sweeps behavior disapproving | 7.82 (7.81) | 8.54 (8.10) | 9.53 (7.92) | 7.42 (8.17) | 4.63 (5.29) |
| Proportion of sweeps in instruction | 31.47 (16.11) | 30.03 (15.25) | 23.52 (10.99) | 36.79 (19.01) | 37.75 (14.96) |
| Proportion of sweeps in personal care | 12.48 (10.19) | 12.89 (10.20) | 13.38 (10.49) | 13.60 (11.75) | 9.25 (6.93) |
| Instructional level | | | | | |
| Average instructional level | 1.65 (0.33) | 1.68 (0.33) | 1.55 (0.31) | 1.65 (0.34) | 1.69 (0.32) |
| Focus | | | | | |
| Proportion of sweeps in English language arts | 8.84 (8.07) | 9.15 (8.06) | 4.97 (4.59) | 9.61 (8.40) | 11.40 (9.20) |
| Proportion of sweeps in math | 3.63 (5.37) | 2.92 (4.57) | 2.37 (3.24) | 5.32 (7.50) | 4.84 (5.57) |
| Proportion of sweeps in science | 5.32 (7.13) | 5.07 (7.41) | 3.52 (4.45) | 7.89 (9.27) | 5.04 (5.08) |
| Tone | | | | | |
| Average tone | 3.41 (0.32) | 3.42 (0.31) | 3.24 (0.22) | 3.47 (0.37) | 3.51 (0.31) |

*Note.* Standard deviations are in parentheses. CCC = community-based center care; HS = Head Start; FCC = family child care; PSP = public school prekindergarten.

TABLE A5

*Multilevel Models Predicting Process Features From Structural Features, Using Complete Case Analysis*

| | Language use | | Schedule | | | Task | | | | Instructional level | Focus | | | Tone |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Feature | (1) Sweeps listening to children | (2) Sweeps talking to children | (3) Sweeps in whole group | (4) Sweeps in transition | (5) Sweeps in centers | (6) Sweeps behavior approving | (7) Sweeps behavior disapproving | (8) Sweeps in instruction | (9) Sweeps in personal care activities | (10) Average instructional level | (11) Sweeps in English language arts | (12) Sweeps in math | (13) Sweeps in science | (14) Average tone |
| Years of experience | −0.02 (0.03) | 0.13* (0.06) | −0.03 (0.08) | 0.00 (0.05) | 0.04 (0.09) | −0.01 (0.02) | 0.00 (0.04) | −0.01 (0.07) | 0.02 (0.04) | 0.0 (0.00) | 0.02 (0.04) | −0.03 (0.02) | 0.07* (0.03) | −0.00 (0.00) |
| Years of education | −0.28 (0.19) | 0.54 (0.38) | 0.23 (0.49) | 0.25 (0.28) | 0.56 (0.57) | −0.14 (0.11) | −0.39 (0.22) | 0.04 (0.43) | 0.14 (0.27) | −0.02* (0.01) | −0.15 (0.22) | −0.06 (0.14) | −0.05 (0.20) | 0.00 (0.01) |
| Group size | 0.06 (0.10) | −1.70*** (0.21) | −0.27 (0.27) | 0.17 (0.16) | −0.12 (0.32) | −0.02 (0.06) | −0.03 (0.12) | −1.14*** (0.24) | −0.09 (0.15) | 0.00 (0.00) | −0.31* (0.12) | −0.16 (0.08) | −0.07 (0.11) | −0.01 (0.00) |
| Child-to-adult ratio | −0.02 (0.18) | 3.05*** (0.37) | 1.31** (0.47) | −0.07 (0.27) | −0.43 (0.55) | 0.10 (0.10) | 0.40 (0.21) | 1.78*** (0.42) | −0.37 (0.26) | 0.01 (0.01) | 1.05*** (0.21) | 0.22 (0.14) | −0.10 (0.19) | 0.01 (0.01) |
| Uses a formal curriculum | 0.20 (0.61) | −3.31** (1.25) | 0.40 (1.61) | 0.01 (0.93) | −3.19 (1.87) | 0.56 (0.36) | −0.31 (0.72) | −0.31 (1.43) | −1.71 (0.90) | −0.01 (0.03) | 0.13 (0.73) | 0.01 (0.47) | 0.35 (0.66) | −0.02 (0.03) |
| Uses a social-emotional curriculum | 0.13 (0.73) | 0.69 (1.52) | 4.66* (1.91) | 1.53 (1.12) | −5.06* (2.23) | 0.29 (0.44) | 0.60 (0.88) | −0.39 (1.73) | −0.61 (1.08) | 0.04 (0.04) | −0.66 (0.89) | 0.28 (0.58) | 0.46 (0.79) | 0.01 (0.03) |
| *N* | 480 | 480 | 480 | 480 | 480 | 480 | 480 | 480 | 480 | 476 | 480 | 480 | 480 | 480 |

*Note.* All models include a program-level random intercept. All models include indicator variables for whether the program was a public school prekindergarten, Head Start, or family child care, making community-based child care the reference group. Additional covariates in all models include educator's household income and indicators for whether they were White or spoke English as a primary language. Standard errors are in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE A6

*Proportion of Variance in Process Features Explained by Covariates and Structural Features*

| | Language use | | Schedule | | | Task | | | | Instructional level | Focus | | | Tone |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R^2$ | (1) Sweeps listening to children | (2) Sweeps talking to children | (3) Sweeps in whole group | (4) Sweeps in transition | (5) Sweeps in centers | (6) Sweeps behavior approving | (7) Sweeps behavior disapproving | (8) Sweeps in instruction | (9) Sweeps in personal care activities | (10) Average instructional level | (11) Sweeps in English language arts | (12) Sweeps in math | (13) Sweeps in science | (14) Average tone |
| $R^2$ from model including covariates (educator characteristics + provider type) | .02 | .15 | .04 | .06 | .03 | .07 | .06 | .14 | .03 | .04 | 0.07 | 0.08 | 0.08 | 0.12 |
| $R^2$ from model with covariates + structural features | .03 | .31 | .07 | .07 | .06 | .08 | .08 | .20 | .05 | .07 | 0.12 | 0.09 | 0.10 | 0.13 |

*Note.* $R^2$ comes from regression analyses with complete cases and standard errors clustered at the provider level. An alternative approach to understanding the variance explained by covariates and structural features would be to calculate pseudo-$R^2$ using variance components yielded from multilevel models, as explained by Singer and Willett (2003). However, using this approach sometimes yields pseudo-$R^2$ values less than 0, which are not interpretable (Singer & Willet, 2003). This was the case with several outcomes in our data, and as such, we report only $R^2$.

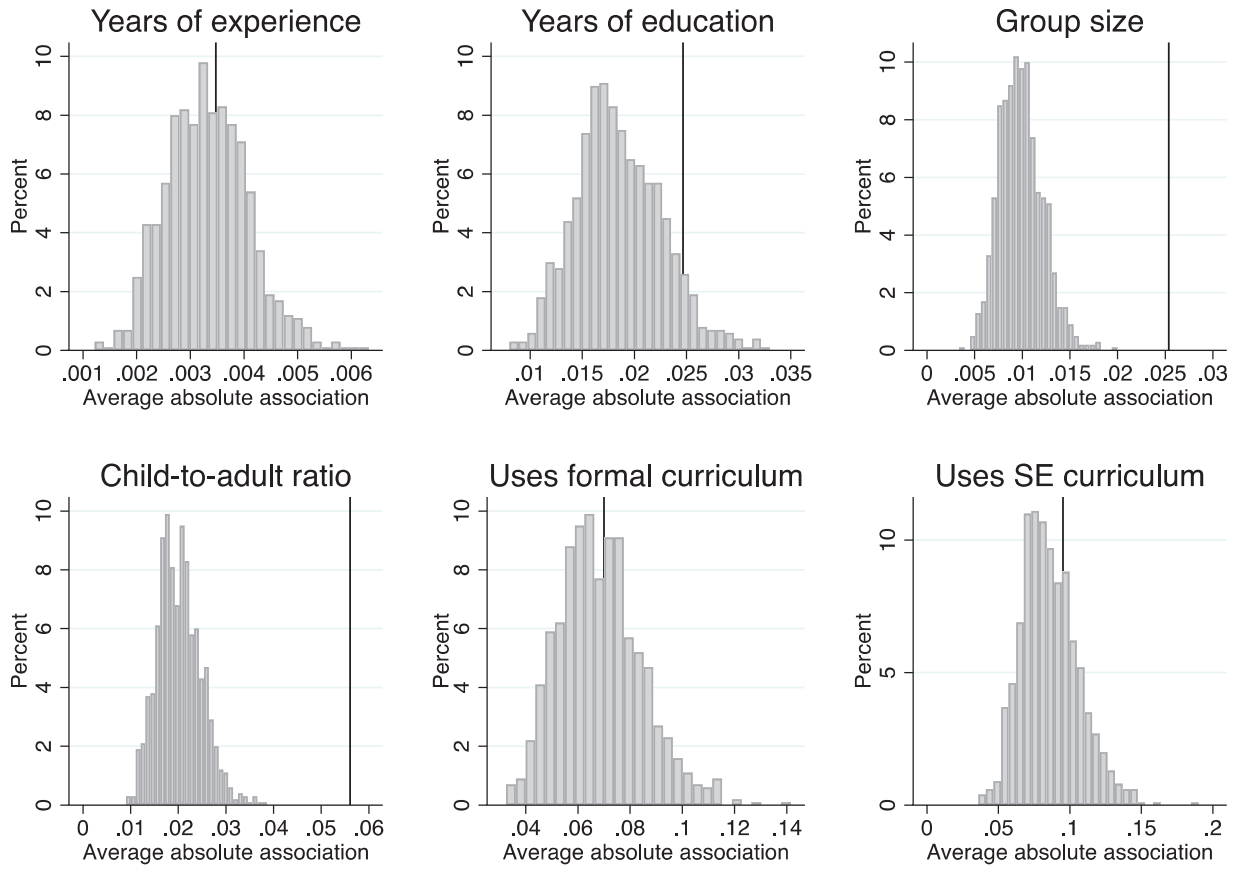


FIGURE A1 *Approximate sampling distribution for the average association between each structural quality feature and process features, using complete case analysis.*

*Note.* The distribution comes from 1,000 trials with randomized structural features using complete-case analysis (i.e., excluding observations missing on any key analytic variables). The vertical line indicates the observed absolute average association between the particular structural feature and the 14 features of process quality in the observed data set. All process features were standardized prior to the analyses using the permutated data. The models controlled for provider type, educator demographics, and other structural features.

## Authors' Note

Analysis files for the current article can be found at https://doi.org/10.3886/E148403V1.


## Acknowledgments

We thank Abt Associates, in particular Amy Checkoway, Barbara Goodson, and Kerry Hofer, for their expertise, ongoing collaboration, and support of this work. We are also deeply grateful to all the educators who have so generously shared their time and perspectives with us. The research activities were funded by the Saul Zaentz Charitable Foundation through its generous support of the Saul Zaentz Early Education Initiative at Harvard Graduate School of Education.


## Notes

1. Recent discussions of classroom observational measures have surfaced grain size as a key dimension differentiating measures. We adopt the terminology of Pianta et al. (2020), who explain that observational measures sit on a continuum from molecular to molar, with more molecular measures capturing the discrete components of processes and more molar measures capturing their abstract aspects.

2. The authors also identified several features of *child* behaviors that they believe are markers of high-quality early learning and care environments that are captured using the Child Observation in Preschools (COP) protocol (Farran & Anthony, 2014). Given that existing measures of process quality, like the CLASS, tend to focus solely on *adult* behaviors, the current article similarly considers only adult behaviors.

3. This implies that most parents likely provided information on the child's lead teacher, although we recognize that in some of the settings observed in these data, the concept of a "lead" is irrelevant as there may have been only one adult or multiple coteachers (e.g., as in many FCC programs).

## References


Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly*, *23*(3), 299–313. https://doi.org/10.1016/j.ecresq.2008.03.001

Bassok, D., Fitzpatrick, M., Greenberg, E., & Loeb, S. (2016). Within- and between-sector quality differences in early childhood education and care. *Child Development*, *87*(5), 1627–1645. https://doi.org/10.1111/cdev.12551

Bilbrey, C., Vorhaus, E., & Farran, D. C. (2007). *Teacher observation in preschool*. Peabody Research Institute, Vanderbilt University.

Build Initiative & Child Trends. (2019). *Quality Rating and Improvement Systems compendium* [Data system]. http://qualitycompendium.org/

Burchinal, M., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*(1), 2–11. https://doi.org/10.1207/S1532480XADS0601_01

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, *112*(3), 579–620.

Chaudry, A., Morrissey, T., Weiland, C., & Yoshikawa, H. (2017). *Cradle to kindergarten: A new plan to combat inequality*. Russell Sage Foundation.

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal*, *45*(2), 443–494. https://doi.org/10.3102/0002831207312908

Cryer, D., Tietze, W., Burchinal, M., Leal, T., & Palacios, J. (1999). Predicting process quality from structural quality in preschool programs: A cross-country comparison. *Early Childhood Research Quarterly*, *14*(3), 339–361. https://doi.org/10.1016/S0885-2006(99)00017-4

Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., & Jones, D. (2009). Fostering high-quality teaching with an enriched curriculum and professional development support:The Head Start REDI program. *American Educational Research Journal*, *46*(2), 567–597. https://doi.org/10.3102/0002831208328089

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., Howes, C., & Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, *21*(2), 174–195. https://doi.org/10.1016/j.ecresq.2006.04.004

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., Cai, K., Clifford, R. M., Ebanks, C., Griffin, J. A., Henry, G. T., Howes, C., Iriondo-Perez, J., Jeon, H.-J., Mashburn, A. J., Peisner-Feinberg, E., Pianta, R. C., Vandergrift, N., & Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, *78*(2), 558–580. https://doi.org/10.1111/j.1467-8624.2007.01014.x

Farran, D. C., & Anthony, K. (2014). *Child Observation in Preschools (COP): MNPS edition*. Peabody Resarch Institute, Vanderbilt University.

Farran, D. C., Meador, D., Christopher, C., Nesbitt, K. T., & Bilbrey, L. E. (2017). Data-driven improvement in prekindergarten classrooms: Report from a partnership in an urban district. *Child Development*, *88*(5), 1466–1479. https://doi.org/10.1111/cdev.12906

Fuhs, M. W., Farran, D. C., & Nesbitt, K. T. (2013). Preschool classroom processes as predictors of children's cognitive self-regulation skills development. *School Psychology Quarterly*, *28*(4), 347–359. https://doi.org/10.1037/spq0000031

Hamre, B. (2014). Teachers' daily interactions with children: An essential ingredient in effective early childhood programs. *Child Development Perspectives*, *8*(4), 223–230. https://doi.org/10.1111/cdep.12090

Hamre, B., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 US early childhood and elementary classrooms*. Foundation for Childhood Development.

Hanno, E. C., Gonzalez, K. E., Lebowitz, R. B., McCoy, D. C., Lizarraga, A., & Korder Fort, C. (2020). Structural and process quality features in Peruvian early childhood education settings. *Journal of Applied Developmental Psychology*, *67*, Article 101105. https://doi.org/10.1016/j.appdev.2019.101105

Hanno, E. C., Jones, S. M., & Lesaux, N. K. (2021). *Back to basics: Developmental catalysts of quality improvement in early education and care*. Manuscript submitted for publication.

Harms, T., & Clifford, R. M. (1980). *Early Childhood Environment Rating Scale*. Teachers College Press.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, *23*(1), 27–50. https://doi.org/10.1016/j.ecresq.2007.05.002

Hu, B. Y., Fan, X., Wu, Y., & Yang, N. (2017). Are structural quality indicators associated with preschool process quality in China? An exploration of threshold effects. *Early Childhood Research Quarterly*, *40*, 163–173. https://doi.org/10.1016/j.ecresq.2017.03.006

Hu, B. Y., Mak, M. C. K., Neitzel, J., Li, K., & Fan, X. (2016). Predictors of Chinese early childhood program quality: Implications for policies. *Children and Youth Services Review*, *70*, 152–162. https://doi.org/10.1016/j.childyouth.2016.09.013

Jenkins, J. M., Whitaker, A. A., Nguyen, T., & Yu, W. (2019). Distinctions without a difference? Preschool curricula and children's development. *Journal of Research on Educational Effectiveness*, *12*(3), 514–549. https://doi.org/10.1080/19345747.2019.1631420

Jones, S. M., Lesaux, N. K., Gonzalez, K. E., Hanno, E. C., & Guzman, R. (2020). Exploring the role of quality in a population study of early education and care. *Early Childhood Research Quarterly*, *53*, 551–570. https://doi.org/10.1016/j.ecresq.2020.06.005

Justice, L. M., Mashburn, A. J., Hamre, B., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51–68. https://doi.org/10.1016/j.ecresq.2007.09.004

Lin, Y.-C., & Magnuson, K. A. (2018). Classroom quality and children's academic skills in child care centers: Understanding the role of teacher qualifications. *Early Childhood Research Quarterly*, *42*, 215–227. https://doi.org/10.1016/j.ecresq.2017.10.003

Markowitz, A. J., Bassok, D., & Hamre, B. (2017). Leveraging developmental insights to improve early childhood education. *Child Development Perspectives*, *12*(2), 87–92. https://doi.org/10.1111/cdep.12266

Mashburn, A. J., Pianta, R. C., Hamre, B., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x

National Center for Early Childhood Quality Assurance. (2017). *Indicators of quality for ratings*. https://childcareta.acf.hhs.gov/sites/default/files/public/qris_indicators_quality_2016.pdf

NICHD Early Child Care Research Network. (2002). Child-care structure → process → outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, *13*(3), 199–206. https://doi.org/10.1111/1467-9280.00438

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott–Shim, M. (2000). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly*, *15*(4), 475–496. https://doi.org/10.1016/S0885-2006(01)00077-1

Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, *12*(3), 281–303. https://doi.org/10.1016/S0885-2006(97)90004-1

Pianta, R. C., Hamre, B. K., & Nguyen, T. (2020). Measuring and improving quality in early care and education. *Early Childhood Research Quarterly*, *51*, 285–287. https://doi.org/10.1016/j.ecresq.2019.10.013

Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. https://doi.org/10.1207/s1532480xads0903_2

Pianta, R. C., La Paro, K., & Hamre, B. (2008). *Classroom Assessment Scoring System (CLASS) manual, pre-K*. Paul H. Brookes.

Sheridan, S., Williams, P., & Samuelsson, I. P. (2014). Group size and organisational conditions for children's learning in preschool: A teacher perspective. *Educational Research*, *56*(4), 379–397. https://doi.org/10.1080/00131881.2014.965562

Sherman, R. A., & Funder, D. C. (2009). Evaluating correlations in studies of personality and behavior: Beyond the number of significant findings to be expected by chance. *Journal of Research in Personality*, *43*(6), 1053–1063. https://doi.org/10.1016/j.jrp.2009.05.010

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurence*. Oxford University Press.

Slot, P. L. (2018). *Structural characteristics and process quality in early childhood education and care: A literature review* (OECD Education Working Papers No. 176). https://doi.org/10.1787/edaf3793-en

Slot, P. L., Bleses, D., Justice, L. M., Markussen-Brown, J., & Højen, A. (2018). Structural and process quality of Danish preschools: Direct and indirect associations with children's growth in language and preliteracy skills. *Early Education and Development*, *29*(4), 581–602. https://doi.org/10.1080/10409289.2018.1452494

Slot, P. L., Leseman, P. P. M., Verhagen, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childhood Research Quarterly*, *33*, 64–76. https://doi.org/10.1016/j.ecresq.2015.06.001

StataCorp. (2021). *Stata Statistical Software: Release 17*. https://www.stata.com/products/

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*, *98*(1), 63–74. https://doi.org/10.1037/0022-0663.98.1.63

Weiland, C., McCormick, M., Mattera, S., Maier, M., & Morris, P. (2018). Preschool curricula and professional development features for getting to high-quality implementation at scale: A comparative review across five trials. *AERA Open*, *4*(1). https://doi.org/10.1177/2332858418757735

Whitebook, M., Phillips, D., & Howes, C. (2014). *Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the National Child Care Staffing Study*. Center for the Study of Child Care Employment.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., Ludwig, J., Magnuson, K. A., Phillips, D., & Zaslow, M. J. (2013, October 11). *Investing in our future: The evidence base on preschool education*. Society for Research in Child Development and Foundation for Child Development. http://fcd-us.org/resources/evidence-base-preschool

## Authors

EMILY C. HANNO is a postdoctoral researcher at the Harvard Graduate School of Education.

KATHRYN E. GONZALEZ is a researcher at Mathematica Policy Research and former graduate student at the Harvard Graduate School of Education.

STEPHANIE M. JONES is the Gerald S. Lesser Professor in Early Childhood Development and codirector of the Saul Zaentz Early Education Initiative at the Harvard Graduate School of Education.

NONIE K. LESAUX is the Juliana W. and William Foss Thompson Professor of Education and Society and codirector of the Saul Zaentz Early Education Initiative at the Harvard Graduate School of Education.