Where has all the Time Gone?

Describing Time Use in Full- vs. Half-day Pre-Kindergarten

Hannah Denker, doctoral candidate
University of Colorado-Boulder School of Education
hannah.denker@colorado.edu
Campus Box 249
Boulder, CO 80309-0249

Allison Atteberry, associate professor
University of Virginia - Curry School of Education
aa2pn@virginia.edu
417 Emmet St S
Charlottesville, VA 22903

Support for this work has also been provided by the Laura and John Arnold Foundation and the Smith Richardson Foundation. The views expressed in the paper are solely those of the authors. Any errors are attributable to the authors.

Abstract

High-quality, early childhood education (ECE) experiences can improve children's learning trajectories and produce long-term social returns. Yet, there is also evidence that ECE effects are variable and do not always persist. There is therefore a need to deepen the research base on what constitutes a 'high quality' ECE experience. This need exists on at least two levels—first, to identify which structural features reliably deliver effects, and second to capture the in-classroom processes those structures make possible. This paper provides a case study using an intensive in-classroom data collection effort—a total of 114 all-day class observations with every minute coded for content area(s) and activity type, and up to 14 repeated observations per teacher over a two-year period. In this instance, the data allow us to gain important insights into how teachers leverage time differently when one structural feature of preschool programming—length of the preschool day—changes.

Keywords: early childhood, educational policy, time use; descriptive analysis

Introduction

High-quality, early childhood education (ECE) experiences can have profound positive effects on child development and yield substantial social returns (Blau & Currie, 2006; Deming, 2009; Heckman, 2006; Nores et al., 2005; Shonkoff et al., 2000; Weiland & Yoshikawa, 2013; Wong et al., 2008). Further, ECE access may contribute to narrowing outcome inequality, since children from historically underserved populations tend to benefit more from ECE interventions. Despite this promise, there is also concerning evidence that ECE effects are variable and that not all ECE experiences yield lasting effects. (Fitzpatrick, 2008; Li et al., 2020; Lipsey et al., 2018; Magnuson et al., 2007; Puma et al., 2012). A recent consensus statement by a group of ECE researchers refers to the 'black box' problem: While there is compelling evidence that some ECE programming *can* impact children's outcomes, we know much less about what happens inside successful interventions to change learning opportunities and outcomes (Phillips et al., 2017).

There is therefore a need to deepen the research base on what constitutes a 'high quality' ECE experience. This need exists on at least two levels—first, to identify which structural features tend to yield larger effects, and second to capture the in-classroom processes those structures make possible. The current paper provides a case study of an intensive, in-classroom, data collection effort that may be useful to future researchers tackling this 'black box' question. In this instance, the data allow us to gain important insights into how teachers leverage time differently when one structural feature of preschool programming—length of the preschool day—changes.

Current Paper within Context of Larger Study

The data used in the current, observational study were collected as part of a broader randomized control trial (RCT) on the causal effects of full- vs. half-day pre-kindergarten (pre-K). The larger study is referred to as the Full-Day Pre-K (FDPK) Study and took place in Westminster

Public Schools (WPS), a predominantly Hispanic and low-income school district near Denver. The child-level randomization is not directly relevant for the current paper's descriptive analyses, as they are conducted at the classroom level. However, we briefly introduce the larger Study and its initial findings to set the stage for examining ECE time use. The broader FDPK Study has two main aims: First, it estimates experimental impacts of full- vs. half-day pre-K offers on a variety of short- and longer-term child and family outcomes. In fact, published findings from the initial cohort show that randomized offers of full- vs. half day spots caused a 0.27 standard deviation (SD) increase on an end of pre-K measure of receptive vocabulary, and similarly large effects on 5 of the 6 domains included in a teacher-administered, observational assessment (Author(s)).[1]

This early evidence of short-term effects underscores the importance of the FDPK Study's second aim: To provide insights into exactly how ECE class time was used.[2] To achieve this aim, the research team executed a multimodal, intensive, data collection effort that gathered information about students' classrooms and teachers. This data collection included surveying teachers about their time use as well as observing and systematically coding the activity and content focus of every FDPK Study teacher's classroom. Thus, these data were collected up to 14 times across 2 years of the study and contain minute-by-minute records for an entire class period. This uniquely dense data collection enables us to share with researchers and policymakers exactly what teachers did with their extra time in this full-day pre-K setting.

In the current paper, we leverage the FDPK Study's class-level observational data to make three contributions to the literature on ECE time use. The first research opportunity arises as a result of how time-use data were collected: The research team initially elected to capture typical time use using an annual teacher self-report survey, which is both relatively easy to collect and is the most common approach in prior large studies.[3] However, the research team was particularly

concerned about providing an *accurate* description of time-use. For the second and third pre-K cohorts, we therefore supplemented the teacher surveys with having trained observers record minute-by-minute time use in each class using an observational protocol. In this paper, we focus on the second and third pre-K cohorts, wherein these dual approaches allow us to consider whether teachers surveys, while much simpler to collect, can provide accurate snapshots of time allocation.

Second, showing how time was used differently in full- vs. half-day classes among teachers working in the same district, in the same year, and using the same curriculum represents a novel contribution to the literature on pre-K. Yet, capturing typical ECE class time is not straightforward. Whereas K-8 classroom time use can often be described using subject-specific time blocks organized by the teacher (e.g., the math block), ECE teachers often mix multiple contents together. Describing time use in pre-K is further complicated because time can be coded for its content focus *and* simultaneously for the activity type (e.g., teacher-directed whole group, centers, etc.). Our data allow us to capture these complexities as we describe a typical pre-K day in both the full- and half-day setting. Finally, we also demonstrate how the observational data can reveal meaningful variability in time use both across classes and within-classrooms over time. While the current data were collected as part of the larger RCT on full- vs. half-day pre-K, the research questions in the current paper are descriptive in nature. We address the following descriptive research questions:

1. *How does teacher self-reported time allocation differ from observed time use?*
2. *How does typical class time allocation differ between full- and half-day pre-K?*
3. *How much does instructional content time allocation vary across teachers and over time?*

Competing Hypotheses about Pre-K Time Use

It is perhaps obvious why capturing class time-use is central to understanding children's educational outcomes. The instructional content and activities to which a child is exposed, along

with the quality of classroom interactions, are the essence of a child's educational experience. Yet, there are competing hypotheses about how the additional hours provided by a full-day pre-K experience should be allocated to maximize child outcomes. Some might anticipate that the FDPK Study's positive full-day effects on academic domains would correspond with more time dedicated to core academic subjects like English language arts (ELA) or math. Indeed, K has certainly become more academic over time (Bassok et al., 2016), and there is even evidence that children *entering* K in 2010 are more proficient in both math and literacy than their 1998 counterparts (Bassok & Latham, 2014). There is also evidence—again for K-age children—that additional time spent on advanced math content is positively associated with student learning (Engel et al., 2016).

Yet what may be developmentally appropriate in K is not necessarily the same for children in pre-K. Stipek et al. (1995) finds that 4- to 6-year old children in more didactic ECE programs tend to have more negative outcomes on socio-emotional and motivational measures, and even on some achievement measures. Instead, a more developmental perspective might suggest that organized play and child-directed activities are the more effective ways to leverage extra time (see e.g., Golinkoff et al., 2006). It could also be that some of the operational components of full-day pre-K benefit young children. For instance, the estimated early effects from the FDPK study span cognitive, socio-emotional, and some health outcomes (special needs and gross motor development). This might suggest that something structural about the provision of full-day care supports development, such as regulating child nap/sleep cycles or providing nutritional lunches.

Literature on Class Time Use for Younger Children

Time Allocation in K and Pre-K

Before we can ask questions about how to *best* allocate classroom time, we must consider the extent to which we know what typically happens during the pre-K day. More time-use research

has focused on K, which may help shape our expectations for pre-K. The nationally representative Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) data provides some insight into class time, though unfortunately its teacher survey does not capture proportions of the K day spent on various content or activities. Bassok et al. (2016) finds in the ECLS-K:2011 that about 97% of K teachers report teaching reading and language arts daily and about 91% of teachers report teaching mathematics daily. Other subjects (e.g., social studies and science) are taught less frequently with no more than 24% of teachers reporting daily instruction on non-core subjects.

In New York City Public Schools, Engel et al. (2021) find that K teachers focus on either math or literacy for over one-third of the school day, on average. They also document the intriguing finding that K students spend large proportions of their day (about 40%) on activities that do not involve instruction or have a content area focus. La Paro and her colleagues (2009) find similar results for K students with data from the National Center for Early Development and Learning's (NCEDL) Multi-State Study of Pre-K. Using the Emerging Academics Snapshot (Ritchie et al., 2001) observation protocol, La Paro et al. finds that students are not engaged in any instructional activity for about 39% of the recorded time—the same amount of time spent on math and literacy combined in the Multi-State Study data.

We are aware of four time-use studies that are specific to pre-K. First, in Tulsa Head Start and Tulsa school-based pre-K, Phillips et al. (2009) find that between 18% and 30% of instructional time is devoted to literacy activities—a little less than in the K studies. Second, Piasta et al. (2014) sampled 65 preschool classrooms in Ohio to investigate preschool mathematics and science learning opportunities. These authors find higher estimates of preschool instructional time devoted to mathematics and science (about 25% of total instructional time in each domain) than documented in previous pre-K observational data (i.e., Connor et al., 2006; La Paro et al.,

2009). Third, Connor et al. (2006) use videotaped observations of Head Start and state-funded preschool programs operating in the same midwestern, urban school district to identify associations between the frequency of literacy activities and child outcomes in ELA. Using only 60- or 120-minute snapshots of the school day, Connor et al. document that these pre-K classes spend between 12 and 14 minutes per day on each of the following: ELA, transitions, art, and receiving instructions. These classes were also observed spending between 2 and 4 minutes per day on math, science, music, and play.

Finally, several researchers use the Multi-State Study data and its follow-up study, the State-Wide Early Education Programs Study (SWEEP),[4] to describe classroom structures, including time use, in pre-K (Chien et al., 2010; Early et al., 2005, 2010; La Paro et al., 2009; Pianta et al., 2005). First, these data also indicate that a substantial part of the pre-K day lacks instructional content (i.e., basics/routines). Second, children spend 20% to 30% of observed time on math and literacy—again, a little less than in K. Finally, NCEDL data shed light on an important an aspect of class time-use for younger children: Pre-K teachers will likely spend less time in the more teacher-centered, single-content activities that become so pervasive once formal schooling begins. Indeed, while reports vary somewhat across studies, researchers using NCEDL data generally find that children spend 24% to 34% of the day in arranged in free choice/centers, as well as in whole group activities (Chien et al., 2010; Early et al., 2005, 2010; La Paro et al., 2009; Pianta et al., 2005).

Complexities of Capturing Time Use: Multidimensionality

As described above, many classroom time-use studies describe time allocation largely based on time dedicated to a *primary* content area. Yet in pre-K, it may be less developmentally appropriate to expect children to sit in one place while their teacher delivers clearly delineated

blocks of content-specific instruction. In fact, one of the core considerations listed in the National Association for the Education of Young Children (NAEYC)'s Developmentally Appropriate Practice (2020) position statement suggests that an integrated curricular approach is more meaningful than teaching content areas in isolation. ECE teachers may weave multiple content areas into more child-directed, play-based activities as recommended by NAEYC. Take, as an illustrative example, a block of time in a pre-K classroom where children are allowed to select one of several learning stations set up throughout the room. Imagine that at one of these stations, or centers, children practice writing their names on paper and then decorating it with different art mediums. In terms of content, the children are learning how to write (ELA) and are making artwork (art/music). In terms of activity organization, the class is divided up into centers. One could categorize time at the writing center in terms of the content children are practicing, the activity grouping the teacher has chosen for the class, or by the child-directed nature of this particular activity.

One of the ways researchers have tackled this complexity in ECE time allocation is to try to identify a set of distinct classroom "profiles" that correspond with a particular mix of activity opportunities. For example, Fuligni et al. (2012) uses observational data to classify classrooms into two different groups—the High Free-Choice and Structured-Balanced profiles. The authors find that the majority of the pre-K day across classrooms is spent in the free-choice activity setting (40%). Chien, et al. (2010) also develop classroom profiles from a latent class analysis. They report that children spend the largest proportion of their day in free-choice and whole-group activities (30% and 27%, respectively). Taken together, this evidence suggests that variation in time use exists at a classroom-level and is likely influenced by a classroom teacher's instructional preferences.

Time Use Measurement Tools: Lessons in Accurately Measuring Classroom Time Use

Classroom time use is typically measured with one of three different instrument types—external observational records, surveys, or time-use logs. However, each of these tends to balance the ease/difficulty of data collection with the quality of the data they yield. Table 1 provides a summary of findings across studies that compare classroom time-use descriptives across at least two instruments. In order to examine the accuracy of self-reported data, several studies compare survey or log data to classroom observations (Burstein, 1995; Camburn et al., 2001; Camburn & Barnes, 2004; Elias & Wheeler, 1976; Mayer, 1999; Rowan & Correnti, 2009; Smithson & Porter, 1994). For example, Camburn and his coauthors assess the validity of elementary teacher log data, relative to classroom observations for documenting reading instructional practices (Camburn et al., 2001; Camburn & Barnes, 2004). These studies find that teachers report more incidences of reading comprehension activities than did classroom observers, and they note agreement among the ELA activities identified by teachers in their logs and independent raters occurred about 73% of the time (Camburn & Barnes, 2004, p. 56). As shown in Table 1 the agreement rates between teacher log data and classroom observations observed in previous literature range from 35% to 96%.

[Insert Table 1 about here]

Other researchers compare teacher logs, an instrument that is typically designed to capture more nuance and complexity with frequent administration, to teacher surveys. For example, Burstein et al. (1995) compare data collected from surveys and daily logs over a five-week period. Importantly, Burstein et al. find, that for all 13 items asking about instructional activities, the rate of direct agreement is never higher than 60%. Especially pertinent to the current study, Rowan and

Correnti (2009) use a hierarchical linear model to document that "large variation…exists in teaching practices—even among teachers working at the same grade in the same school" (p. 126).

In sum, previous literature quantifies the extent to which instruments can accurately measure structural features of a classroom. Together the evidence suggests that, when it is feasible, classroom observations provide the most direct measures of classroom practices. However, since classroom observations are costly and time-consuming, researchers may still opt for the much more pragmatic approach of surveying teachers about their own time use.

Contribution to the Literature

We build on the existing literature in a few ways. The literature suggests that we should expect that pre-K teachers will spend the majority of their instructional time teaching reading and math, but that a large part of their day will be spent in activities with no instructional content (Early et al., 2010; Engel et al., 2021). Like the Emerging Academics Snapshot (Ritchie et al., 2001) observation protocol used in the NCEDL studies, our observation protocol enables us to acknowledge the complexities of time allocation in classes with young children by both coding for multiple content areas taught simultaneously (e.g., literacy and social studies) and also coding for activity type. We delve into the mixed content question by reporting which content areas are most commonly mixed together in our pre-K setting. We also describe typical time use in two ways—first in terms of content focus and then by activity type.

We also contribute to the literature a comparison of the full- vs. half-day pre-K settings to consider whether a longer school day potentially affords additional opportunity to take up content and activities in more varied ways. We also extend prior research, which has focused on describing an average or typical day, by focusing on variation in time allocation across pre-K teachers and also within the same teacher's classroom over time.

WPS and its Pre-K Offerings

All of the FDPK Study classrooms were located within the public-school district of WPS. Located near Denver, this school district is typically-sized, serving about 10,000 students annually. In 2015-16, the district's student population was 83% eligible for the federal free/reduced price lunch (FRPL) program and 77% Hispanic (U.S. Department of Education, National Center for Education Statistics, 2016). WPS also serves a sizeable English Language Learner population (34%). The WPS demographics make this an especially important context for ECE research, since some studies have shown that ECE especially benefits low-income (Deming, 2009; Magnuson et al., 2007; Weiland & Yoshikawa, 2013) and Hispanic children (Currie & Thomas, 1999; Gormley & Gayer, 2005). Like WPS, thousands of U.S. districts serve a predominantly FRPL-eligible student population. While WPS does serve a larger than average proportion of Hispanic students, 20% of U.S. public school districts are at least 25% Hispanic (Reardon et al., 2021).

Prior to the FDPK Study, WPS only offered half-day pre-K classes (4 days per week for 3 hours per day in an AM or PM session), located primarily at a single school site. In 2016-17, WPS piloted 7 new full-day classes, each one added to an existing elementary school site in the district. Full-day classes met 5 days per week for about 7 hours a day, more than doubling pre-K time from about 510 total hours across the pre-K year to 1190 hours. Due to oversubscription to the new full-day classes, the district opted to use a randomized lottery[5] to fairly allocate these limited spots to eligible families in three consecutive cohorts of pre-K applicants—Cohort 1 for 2016-17, Cohort 2 for 2017-18, and Cohort 3 for 2018-19.[6] The current paper examines observational time-use data, which was only collected for Cohorts 2 and 3. Over this two-year period, we conducted 114 repeated observations of 34 pre-K classes and administered 24 teacher surveys.

A few other aspects of the WPS setting are relevant for framing expectations for pre-K class time-use. First, all ECE classes—AM, PM, and full-day—had approximately 16 students. All teachers taught a full day schedule (half-day teachers taught one class of ~16 children in the morning, and a separate class of ~16 children in the afternoon). All WPS pre-K teachers—both in full-and half-day—used the same curricular resources throughout the FDPK Study.[7] Any full-vs. half-day time-use differences are therefore not an artifact of using different curricula. Second, these curricula are not prescriptive about time allocation (as evidenced by the fact they can be readily adapted by teachers as needed for half- or full-day settings). We therefore do not anticipate that curricula will *dictate* observed time use. It is also worth noting that the mission statement of WPS's early childhood education program describes their program as "play-based…that prepares the whole child for success in elementary school" (Westminster Public Schools, 2021). One may anticipate that teachers in this district would favor group- and play-based instruction across content areas.

We expect that the state's preschool standards could shape WPS teachers' choices in how to allocate the additional time afforded by a full-day schedule. The Colorado Academic Standards, originally adopted in 2009-2010, set standards in 10 content areas for preschool children. If the number of standards within each content area corresponds to the amount of time needed to cover the content, this could guide hypotheses about how we would expect WPS teachers to use their class time. The standards suggest we should expect an emphasis on ELA (33 separate standards) and social studies (23). Math and drama/theater had the fewest standards at 9 and 7, respectively. Consequently, one would expect that teachers would not use as much class time on these two areas.

Data, Analytic Sample, and Measures

Data Collected on Pre-K Time Use

As mentioned, the research team implemented two different approaches to collect information about classroom instruction (for an overview of the data collection timing, see Table 2). The first approach was to administer a survey to pre-K teachers at the end of each school year in each pre-K cohort (92% response rate). We also collected classroom data by having trained observers spend a whole day in each study classroom several times per year. Observers used an adapted version of the *Advanced Narrative Record Observation for Classrooms* developed by Farran and her coauthors (2015) as the observational protocol.[8] The Narrative Record combines observer and instructional notes in an open-ended format to provide an uninterrupted account of everything that occurs during a classroom observation. From the Narrative Record data, we compiled detailed information about the instruction taking place in each classroom, including a record of time students spent on different content areas and activities (more on the observational protocol below). In total, we coded over 260 hours of classroom observations for this study.

[Insert Table 2 about here]

Analytic Sample Sizes and Descriptives

The current paper is limited to the half- and full-day classes in the FDPK Study's second and third cohorts. Since some analyses are conducted at the classroom level but others at the class level, refer to Table 2 for a quick overview of sample sizes.[9] The first column in Table 2 provides the count of unique teacher class*rooms*, separately by cohort. The second column counts the number of unique *classes*—that is, groups of students. Since teachers in half-day classrooms serve both an AM class and a PM class each with distinct sets of students, the number of half-day classes is twice the number of half-day classrooms. While there were 34 unique *classes* of children (18 full-day + 8 AM half-day + 8 PM half-day), there were a total of 26 unique class*rooms* across the two cohorts (18 full-day classrooms + 8 half-day classrooms with both an AM and PM session).

Because many teachers taught in both Cohort 2 and Cohort 3, these 26 classrooms were led by 16 unique teachers (6 half-day and 11 full-day teachers).[10] The relative advantage of the current data is the depth of insight into classrooms, not the breadth across a large number of teachers that would allow for inferences to be made about class-level relationships (e.g., between teacher characteristics and time-use). This paper, therefore, does not attempt to isolate a causal link between full-/half-day class structure and time-use, estimate causal effects, or rule out potential bias in those effects. That said, it is still useful to consider the characteristics of full- vs. half-day teachers (Table 3) before examining our descriptive findings. We might interpret patterns differently if, for instance, full-day teachers were always more experienced, had more certifications, and had higher levels of education.

While some of the percentage point differences in Table 3 may at first seem large, the absolute differences are small (e.g., 50% vs. 90% of half- vs. full-day teachers are white, but this only represents a difference of 3 vs. 9 teachers). In general, we do not see clearly systematic patterns across the characteristics of half- vs. full-day teachers in Table 3. Full-day teachers are slightly less experienced but are more likely to have a BA or beyond.[11] All teachers are certified. Unsurprisingly, none of these differences in teacher characteristics are statistically significant.

[Insert Table 3 about here]

The study classrooms are spread across 10 total school sites. As noted above, the district's original half-day offerings were located almost entirely at a single school site. When WPS piloted new full-day classrooms, they were each added to a different elementary school. Since WPS is geographically small and has a large majority of both Hispanic and low-income K-12 students in its overall student population, there is limited heterogeneity in student demographics across WPS schools. In Online Appendix A, we present the demographic composition of WPS's pre-K students

in each of study's 10 school sites, alongside the number of full- or half-day classes in each. The percentage of pre-K children identified as Hispanic ranges from 59% to 75% across the 10 school sites. The percentage of children who are FRPL-eligible ranges from 55% to 75%.

Observation- and Survey-Based Time-Use Measures

Observed Time Use

We provide an example of a completed time-use protocol in Online Appendix B. In Cohort 2, the protocol was applied by trained observers in-person, while in Cohort 3 the observers applied the protocol to high-quality video recordings of classrooms.[12] As shown in Table 2, each Cohort 2 classroom was observed up to 3 times per year (Nov/Dec, Feb/Mar, and Apr/May), and each Cohort 3 classroom was observed up to 4 times per year (with an additional Jan/Feb window).

In order to fill out the time-use protocol, observers captured "time-use episodes" throughout the pre-K day. Logged time-use episodes serve as a record of events that happen in the classroom. New time-use episodes begin when the activity type or content type changes and lasts more than a minute. A single time-use episode is continued until 75% of the students have started a new activity or are engaged in new content.

Time is coded in several ways during the classroom observations, two of which are central to our analyses. First, observers identified the type of activity[13] that was taking place during the episode and, second, they identified between 0 and 3 different content areas[14] being targeted within the given episode. Thus, time was simultaneously coded for activity and content area(s) in the observation protocol.[15] In episodes where activities were not accompanied by any content area instruction (e.g., napping), content was coded as either "no content observed" or "not applicable." Some content/activities coded separately in the time-use protocol were combined to make categories of time use as similar as possible to the survey we administered to the pre-K teachers.[16]

See Online Appendix C for details about how instructional categories were combined and/or restricted. For episodes in which observers identified more than one content area being taught simultaneously, we initially code these as "mixed" content. We subsequently disaggregate mixed content episodes into their primary and secondary content area combinations in order to document which content areas teachers most frequently mixed together.

In Table 4, we present a summary of the number of surveys/observations collected per teacher (N=16 teachers total) across the two cohorts. For instance, Teacher D only appeared in Cohort 2. She taught half-day (AM & PM), completed a survey that year, and was observed 6 total times (3 for AM, 3 for PM). Teacher J, on the other hand, taught full-day pre-K in both cohorts. She filled out a survey both years and was observed a total of 7 times. Teacher J's full-day class was observed three times in Cohort 2 and four times in Cohort 3. Overall in Cohort 2, we observed a total of 13 teachers on up to 3 different days over the course of the school year, resulting in 47 total observations across pre-K classrooms out of a possible 51 (a 92% Cohort 2 coverage rate). In Cohort 3, we observed a total of 13 teachers on up to 4 different days throughout the school year, resulting in 67 total observations across pre-K classrooms out of a possible 68 (a 99% coverage rate). In total, we collected 114 (of 119 possible) class-year observations (96%) across both cohorts.

[Insert Table 4 about here]

Teacher Survey

First, recall that the teacher survey data is only used for RQ 1 (comparing time use across instruments). On the teacher surveys, pre-K teachers were asked to "indicate the number of minutes spent" on a typical day in the following categories: math, reading and language arts, social studies or science, eating, napping or quiet reading, art/music/drama/dance, unstructured play (child-

directed play including recess), structured play (teacher-directed play), and transitioning between activities. We asked about each day of the week separately in order to give teachers the opportunity to report variability that might exist within their weekly schedules, though this rarely occurred.[17] Over the two study years, our research team made minor adjustments to the time-use survey items in order to improve data quality.[18] In both cohorts, 12 of the 13 teachers completed the survey (a 92% response rate) with a different teacher failing to respond each year.[19]

Crosswalk between Survey- and Observation-Based Time-Use Codes

The observational protocol and the teacher survey were developed separately; therefore, they did not use the exact same set of categories to describe class content/activities. However, we are interested in examining how (in)consistent time allocation appears across these two approaches, which requires a crosswalk between the two approaches. In many cases, it was straightforward to equate survey categories with observational categories (e.g., "Napping or quiet resting" from the survey was equated with the observational protocol's "Napping" activity code). However, some substantive analytic choices were required in order to produce a crosswalk between content/activities presented to teachers on the survey with those contained in the Narrative Record observational tool. See Table 5 for a summary of the final crosswalk. For a complete discussion of how these choices were made, see Online Appendix E.

[Insert Table 5 about here]

Results

RQ 1: How Does Teacher Self-Reported Time Allocation Differ from Observed Time Use?

Before characterizing time-use in WPS, we begin with a more methodological question: Did teacher surveys capture accurate insights into class time use? To explain our analytic approach, we use time spent on math as an example.[20] We first create a dataset that reports the number of

minutes spent on math in each observation window, for each class-year, and include the teacher's annual survey (i.e., the dataset is long by observation window, class, year, and data-source). We then employ the following regression model:

$$DailyMinsContent_{swcy} = \beta_0 + \beta_1(Survey_s) + \lambda_{cy} + \varepsilon_{swcy} \quad \text{(Eq. 1)}$$

In Equation (1) the outcome of interest, $DailyMinsContent_{swcy}$, represents the number of minutes spent on a given content area in observation window *w*, in class c, in year y, according to source s (either the survey or the observational data). The predictor variable, $Survey_s$, is an indicator variable that equals 1 if the source of the given outcome is the survey and 0 if it comes from the observational data. To ensure we only examine associations between the 3 or 4 observed results and survey results that belong to the same class in the same year, we control for class-year (i.e., we include a set of class-by-year fixed effects, λ_{cy}). The key parameter of interest is β_1, and it captures how many more/fewer minutes per day—according to the *survey*—teachers reported spending on each content area for the same class in the same year, relative to the average number of minutes actually *observed*. We run the analysis separately by subject and by class type.[21]

In Table 6, we present the mean number of observed minutes in column 1, and in column 2 we report estimates of β_1 (how many more/fewer minutes according to the survey) for each subject. Throughout Table 6, positive estimates of β_1 indicate that teachers self-reported spending more time on a given subject than was actually observed in their class. Negative values indicate time that was underreported on the survey. Take the first row of Table 6 as an example of how to interpret the results: The average number of minutes spent on ELA that *observers* recorded in a given half-day classroom is 30.7 minutes, but we find that the *teachers'* reports differed, on average, by about 11.6 fewer daily minutes.

[Insert Table 6 about here]

Teachers tend to *over*-estimate the amount of time they spend on math, unstructured play, and eating (though only differences in eating are statistically significant). Both half- and full-day teachers tend to significantly *under*-estimate the amount of time they spent on ELA, science/social studies, music/art, and structured play (a difference of between 9 and 19 minutes). In addition, both half- and full-day teachers provide the greatest underestimates of time spent on transitions, with half-day underreporting this time by about 22.1 minutes per day, and full-day teachers underreporting by about 44.6 minutes per day. This suggests that there are clear discrepancies, within many parts of the pre-K day, between how teachers recall typical time use and how much time is actually allocated. Overall, there are statistically significant differences between observations and teacher self-reports across most content and activities, with the exceptions of math and unstructured play for half-day, and ELA, math, unstructured play, and napping for full-day classrooms. Importantly, we see large, statistically significant differences between the two data sources for time spent transitioning.

We also correlate observed time with the teacher-reported time for math and ELA, by half- and full-day (results are shown in the top two rows of Table 1, alongside similar statistics from prior studies). We find that teacher survey data and classroom observation data are positively correlated for math in full-day classes ($\rho = .49$), although this correlation is weaker for half-day classrooms ($\rho = .17$) Our correlations for ELA are less straightforward. We find that the teacher survey and classroom observation data are substantively unrelated in full-day classes ($\rho = -.04$) and exhibit a moderately negative correlation in half-day classes ($\rho = -.33$).

In sum, while teacher surveys are appealing for their relative ease of administration, our results call into question whether they reflect what a trained observer would capture if they spent time in pre-K classrooms throughout the school year. This may be especially true for content areas

or activities that are taught in less structured ways for younger children. For these reasons, we opt to describe time-use going forward using the observational data instead of teacher survey data.

RQ 2: How Does Class Time Allocation Differ between Full- and Half-day Pre-K?

Averaging across the 34 unique class-years, we illustrate the typical half- and full- day by presenting the mean number of minutes spent daily on each of the activity types (Figure 1), and then recoded in terms of content areas (Figure 2). In Table 7, we present these same results but report them as a percent of the school day (mean, SD, 25th and 75th percentile), first by activity type (upper panel) and then again by content area (lower panel). Since we will want to consider time use both proportionally and in an absolute sense, Figure 1 reports time use both as a percent of the day (in bar labels) *and* as minutes spent on each activity type (y-axis).

[Insert Figure 1 and Table 7 about here]

Activity Types

Looking at Figure 1 and the upper panel of Table 7, the most immediate difference in activity time use across half- and full-day classrooms is that full-day students take naps (on average, 71 minutes or 19% of the full-day) while half-day students do not.[22] This is expected, because the school district only requires full-day classrooms to set aside an hour per day for napping or quiet resting.

Aside from napping, half-day classes are similar, proportionally, to full-day classes. Indeed, the proportions of time spent in most activities are all within a few percentage points of one another (the exceptions being whole group activities (27% vs. 17%) and small group plus centers (10% vs. 3%)). Of course, since the full-day is twice as long as the half-day class, we would expect students in full-day to spend around double the number of *minutes* in these activities. This is the case for eating, transitions, small-group only, and centers only activities.

2A. How Much Time is Spent on Non-Instructional Activities?

Figure 1 also illustrates that a sizeable portion of the WPS school day is spent on non-instructional activities (eating, transitions, and playground time). Children in half-day typically spend about 41% of their day—about 76 minutes—on these non-instructional activities. Children in full-day spend a similar proportion of their time on those activities (38%), but with the addition of full-day's nap time, the proportion of time spent in non-instructional activities in full-day is quite high—57% or about 212 minutes per day.

2B. Does Doubling the School Day Double Instructional Activities?

One hypothesis about how to maximize child outcomes with the extra hours afforded by full-day would be to dedicate them primarily to academic instruction. However, the observations reveal that full-day teachers clearly did not simply double the time spent in activities associated with instruction. That said, we will see that the increases they did make, which may seem modest on a day-to-day basis, accumulate to meaningful differences.

First, we consider time allocated to whole group activities, which we would expect is most expressly used to deliver instruction. In half-day, teachers dedicate, on average, 51 minutes (27%) of the day to whole group instruction, while full-day teachers only spend slightly more time—62 minutes (17%) of the day in whole group activities. This finding is interesting because although full-day teachers have about 3 more hours in their school day than half-day teachers (or 2 hours if we exclude napping), they are only spending an average of about 11 more minutes per day in a whole-group setting.

In addition to whole group activities, we would expect to see instruction taking place in the following three "group-based" activity settings: small groups only, centers only, and small groups + centers. We observe that half-day teachers spend an average of 55 minutes (29%) of their day

in these three activities, compared to an average of 86 minutes (23%) in full-day. This amounts to children in full-day classrooms participating in about 30 more minutes per day in group-based activities, relative to their half-day peers. Combining time spent in these group-based activities with time spent in whole group, children in half-day spend about 106 minutes (56%) of the day on activities in which we would expect to see instruction. Had full-day teachers been truly doubling the time spent in instructional settings, we would expect to see an average of 212 minutes per day. In practice, full-day classes spend only about 148 minutes per day on these 4 instructional activities—far from a double dose.

In sum, when we examine the school day through the lens of activity types, we find that the additional 3 hours available in the full-day setting are used first to add a nap and then to spend 10 to 20 additional minutes on each of the activities taking place in the half-day setting. This suggests the additional time is not used to dramatically alter the proportion of time spent on any particular activity (other than napping), but rather to increase time spent on class activities in a somewhat proportional manner. At first, this may not seem like a particularly 'intensive' use of the extra hours in full-day classes. Yet the absolute differences add up across the week. In total, full-day students spend on average 5.3 additional hours each week on these 4 instructional activities, relative to their half-day counterparts.

Content Focus

We also analyze pre-K classroom observations by identifying the primary content focus of time-use episodes. In Figure 2 and Table 7, time with more than one simultaneous content focus is initially coded as "mixed content".

2C. How Much of the Pre-K Day is Content-Free?

About 30% of awake[23] pre-K time does not involve content delivery—both in half- and full-day. In the typical half-day class, 56 minutes (29%) have no content.[24] Turning to full-day, if we exclude the average 71 minutes per day spent napping (refer back to Figure 1), 101 minutes are without a content focus (33% of the awake full day). If we also include the nap as no-content time, 172 minutes, or nearly half of the day (46%), is without instructional content in full-day.

[Insert Figure 2 about here]

2D. How Much Academic Content Do Pre-K Children Receive?

Turning to class time that *did* have a content focus, the first clear finding is that pre-K teachers quite often engage children in activities with more than one simultaneous content focus. Indeed, "mixed content" is the modal content delivered in both half-day (89 minutes or 48% of the day) and full-day settings (128 minutes or 35% of the day). Since mixing content is so prevalent, we will turn to a more in-depth discussion of mixed content in the following section.

In both half- and full-day classes, teachers dedicate about 10% of the day to single academic content areas delivered on their own (ELA, math, science and social studies). However, the absolute difference in daily time again adds up across the week. Relative to their counterparts in half-day classes, children in full-day classes receive, on average, 73 additional minutes focused exclusively on one of these subjects. As expected, most of this single-content instructional time is dedicated to ELA (between 12 and 16 minutes daily). However, if we fail to look more closely into mixed content, it appears that very little time is dedicated to math in either setting (2 to 3 minutes).

2E. How Often Do Teachers Mix Multiple Content Areas Together, and Which Content Areas?

In WPS pre-K classes, it is much more common to observe teachers weaving multiple content areas together than teaching one content area at a time.[25] For an example, consider students

using paint brushes to make artwork of their names. We would code this example as *mixed content delivery* since this assignment simultaneously serves multiple developmental purposes, including learning the alphabet (ELA) and improving hand-eye coordination (fine motor). In half-day classes, almost half of the school day involves mixed content delivery (about 89 minutes). In full-day, children engage in more mixed content learning in an absolute sense (about 128 minutes), though that represents less time in a proportional sense (about a third of the day).

What content areas were most often mixed together? Table 8 provides descriptive statistics in minutes for the 10 most frequently observed content combinations identified as mixed content. Content combinations are listed in order of the frequency in which they were observed in half-day classrooms. About 92% of mixed content episodes are contained in this top 10 list. The mixed content code is used most frequently to describe time with both ELA and social studies/science content. On average, pre-K teachers devoted about 17 minutes per day to this particular content combination across full and half-day settings.

[Insert Table 8 about here]

When we look more closely into mixed content, we get a very different sense of how much ELA and math content children are receiving. ELA content was included in half of the 10 most frequent mixed content combinations. If we sum the average time devoted to mixed content in which ELA present, we find that half-day teachers spend about 88 minutes per day (and full-day teachers about 76 minutes per day) on mixed content that includes ELA. If we next consider math, it is clear that most math instruction in these pre-K classrooms occurred mixed in with other content areas. Observers rarely recorded teachers delivering *only* math (see Figure 2). However, in half-day classes, math appears to be mixed into about 71 minutes of the day (54 minutes per day for full-day classrooms).

RQ 3: How Much Does Instructional Content Time Vary across Teachers and the Year?

While it is important to describe typical time use in pre-K, we believe an understudied question is how much variability lies underneath typical patterns—both across different teachers and within a teacher's classroom over time. By looking across teachers, we can determine what content areas receive similar coverage across all classrooms and, in contrast, where teachers seem to have more discretion. Moreover, a unique aspect of the current study is that many teachers were observed repeatedly over the two-year period, for an average of 6 class periods (minimum of 3, maximum of 14). Analyzing these within-teacher temporal variations may help researchers think about how they should collect their own time-use data. A lack of time-use continuity across the school year would also complicate how we think about survey items that ask teachers to describe a 'typical week'.

Variation Across Teachers in Average Minutes Spent on Key Content Areas

Teachers in WPS—in both half- and full-day—exhibit a significant amount of discretion in how they allocate the time dedicated to ELA, math, science/social studies, music/art, and gross motor skill development. Figure 3 illustrates this point by depicting the variation across teachers in the average minutes they spend on these 5 content areas (half-day teachers on the left, full-day teachers on the right). The bold "X" indicates the mean for that specific content area. Each dot, labeled with a unique letter, represents a specific teacher's average minutes spent per day on a given subject—that is, averaged across all her (up to 14) observations.

[Insert Figure 3 about here]

3A. What Content Areas Exhibit the Least and Most Variation Across Teachers?

In Figure 3, the content area with both the least average time spent ("X") and the least variation across teachers is gross motor skill development. In contrast, both half- and full-day

teachers spend the most time on ELA content, on average ("X"); yet there is evidence that time dedicated to ELA varies from teacher to teacher. For instance, half-day Teacher F typically spends only 18 minutes daily on ELA content, while half-day Teacher B tends to spend twice as much time on ELA (39 minutes per day). This suggests that, even within the more constrained half-day schedule, children with different teachers may be receiving very different dosages of content. As we might expect with more minutes available for allocation, the absolute range of ELA time is even wider among full-day classes. At the low end, full-day Teacher N spends, on average, only 25 minutes on ELA per day. Full-day Teacher N thus spends less time on ELA content than in the average half-day class. At the high end, full-day Teacher G seems to make ELA a priority, dedicating 74 minutes daily to this content area on average.

In RQ2, we focused on meaningfully large differences in average content dosage when comparing the half- and full-day settings. Yet Figure 3 adds nuance to that story, recognizing there is notable time-use variation within these two groups. A half-day teacher can choose to spend just as much time on a given content area as a full-day teacher. Although, the choice to do so might come at a greater cost because they have less total time to allocate.

3B. Do Teachers Balance Time Constraints Differently?

Since there are only so many minutes in the day, teachers who choose to spend a great deal of time on one content area will have less time to spend on others.[26] Figure 3 illustrates this fact, showing that teachers appear to make different tradeoffs. For instance, half-day Teacher D spends more time than any of her half-day colleagues on 3 of 5 five content areas: gross motor skills (14 daily minutes), music/art (39 daily minutes), and science/social studies (58 daily minutes). To accommodate this, Teacher D spends relatively less time on ELA and math. Another example is full-day Teacher G, who we saw spends more daily time than anyone else on ELA content (74

minutes). To accommodate this, however, Teacher G spends less time on math and science/social studies than even the average half-day teacher. The takeaways from Figure 3 illustrate how researchers can go beyond examining average time use to gain insight into how differently teachers allocate time—not only across half-/full-day comparisons, but even within those class types.

Variation Within Teachers Over Time in Minutes Spent on Key Subjects

We now narrow our focus to the two core content areas of ELA and math in Figure 4 (figures similar to Figure 3 and Figure 4 for other activities/content areas are available in Online Appendix F). We are interested in within-teacher time-use variability for at least two reasons: First, we ask whether children in a given classroom are exposed to relatively consistent time allocations from one day to the next, and we are particularly interested in how this differs for half- vs. full-day classes. Second, these analyses serve as a tool to help future researchers think about the number of classroom observations needed to get a sense of a teacher's 'typical' time allocation.

[Insert Figure 4 about here]

3C. Do Teachers Allocate Consistent Amounts of Time to Core Content Areas Across the Year?

While some teachers allocated a consistent amount of time to ELA or math from one observation to the next, others exhibited a notably wide range. For instance, half-day Teachers D, E, and F are among the more consistent (and consistently low) teachers in terms of daily minutes spent on ELA and math. Full-day Teacher H, on the other hand, did not cover any ELA content during one of her observations; yet, she spent 100 minutes on ELA during another observation. We also can see a tendency for teachers to more consistently spend less time on a given subject, though we do see exceptions to this (e.g., Teacher L exhibits both high and consistent daily minutes on ELA). Finally, as one might anticipate, we see a clear difference between half- and full-day teachers' consistency in time allocation. Half-day teachers have less time each day, and therefore

tend to be more consistent about the amount of time spent on these subjects. Full-day teachers appear to take advantage of the longer day by exhibiting more fluidity in their time use.

3D. How Many Observations are Needed to Capture a Teacher's Typical Time Use?

Since some teachers are clearly more consistent with time allocation across the year than others, it follows that some teachers' time use could be captured accurately with fewer observations. However, unless the researcher already has some sense of how variable teachers' schedules tend to be, the degree of consistency will not be known in advance. Results from this study suggest researchers should anticipate that time use in ECE settings can be highly variable. At a minimum, Figure 4 clearly illustrates that a single observation is unlikely to provide much insight into a teacher's typical time-use patterns. Based on Figure 4, we would suggest at least 3 all-day observations may be required, though different researchers may opt to conduct more observations depending on their tolerance for imprecision due to sampling variability.

Discussion

Study Objectives and Summary of Findings

The current paper provides a case study of how an intensive in-classroom data collection effort allows researchers to capture unique aspects of time allocation in pre-K settings. Over the course of two years, the research team conducted 114 pre-K classroom observations and collected 24 teacher surveys. We address three research questions: First, do teachers' self-report surveys accurately reflect observed time use (RQ1)? Second, how do teachers typically allocate time differently when one structural feature of pre-K programming—length of the pre-K day—changes (RQ2)? Third, to what extent do the typical time-use patterns from RQ2 mask important variability in time-use across teachers and throughout the school year?

Our RQ1 results call into question whether teacher self-reported time use aligns with what trained observers would capture throughout the school year. Teachers tend to underestimate time on certain activities, most notably for time spent transitioning between activities. Half-day teachers underestimated daily transition time by 22 minutes, and full-day teachers underestimated daily transition time by 45 minutes.

For RQ2, we hypothesized several different ways teachers might adapt time-use to the longer school day setting. One approach would be to allot the extra time in the more-than-doubled pre-K school day to increasing the dosage of instructional activities or core content areas (e.g., ELA, math, science, or social studies) by a factor of 2. Another possibility is that children this age would most benefit from learning within longer or more frequent periods of structured play. A third approach might be to do a little more of everything—that is, to proportionally increase the time spent on each of the activities/content are taking place in the half-day setting. Finally, the extra time could be used to engage in brand new activities that were not feasible in a 3-hour window.

We find that, with the exception of a nap period afforded only to students in full-day, the time allocated to most activities increases proportionally in the full-day setting. Our results, therefore, do not support the hypothesis that the doubling of the school-day was used to essentially double the delivery of academic content or time spent in instructional activities. While time-use looks similar across settings in a proportional sense, differences still translate to a sizable, absolute increase in full-day children's exposure to academic content and instructional. Full-day students may spend about 177 additional hours throughout the pre-K year on instructional activities, relative to their half-day counterparts.

We also briefly summarize two other key takeaways for RQ2. First, we find that much of class time in both half- and full-day settings is coded as non-instructional activity types (about 40% of the day) and/or without content (30% in half-day and 46% in full-day). This finding contributes to our body of knowledge on pre-K time use as previous literature has reported anywhere between 11% and 44% of observed instructional time did not contain content (Chien et al., 2010; Coelho et al., 2019; Early et al., 2010; Engel et al., 2021; Fuligni et al., 2012). Our RQ2 findings also highlight that WPS pre-K teachers are much more likely to weave multiple content areas together than deliver single-content blocks. In fact, 48% of observed time in half-day and 35% in full-day classrooms is coded as mixed content. The substantial use of mixed content in WPS pre-K contrasts sharply with what Engel, Jacob, Claessens, and Erickson (2021) find in low-income New York City K classrooms—wherein only 4% of the day is coded as mixed content. Once formal schooling begins, teachers may be required to adopt more traditional content blocks in lieu of mixing content together. In WPS pre-K, however, one would miss much of the content delivery taking place if observers were only allowed to code one content area for any given activity.

For RQ3, we also uncover meaningful variability in average time spent on key content areas across the sample of teachers. We find evidence that there is room for teacher discretion even though half-day teachers are certainly more constrained in how much time they can spend on a given content area. In fact, there are teachers in half-day classes that manage to spend more time on key content areas than their full-day colleagues. These results illustrate how researchers can go beyond examining average time use to gain insight into how differently teachers allocate time—not only across half-/full-day comparisons, but even within those class types.

Finally, our RQ3 work documents how much a teacher's time spent on ELA and math may vary across multiple observations of her class in a two-year period. These results suggest that time may be used very differently from one day to the next in these ECE classrooms. As an illustrative example, take Teacher C, who was observed for the entirety of 8 different (half) days. On one of those days, she spent 0 minutes on math. If we were to only use this observation, we might characterize her as an 'infrequent math' teacher. Yet on another day, Teacher C allocated 100 minutes to math content—one of the highest daily observed times spent on math across the study.

Implications for Research and Policy

Our results may be particularly useful to researchers who study the variation in ECE classroom settings and subsequent consequences on child outcomes. For instance, our RQ1 findings would strongly caution researchers against relying solely on teacher self-reported time-use to describe the pre-K day. These concerns are particularly salient for ECE settings because we also show that describing ECE time-use is complex. Multidimensionality exists both in terms of teachers using a wider range of activity types (e.g., centers, whole group, structured/unstructured play, etc.) and a greater tendency to integrate multiple content areas into a single activity. Second, for researchers who opt to use trained observers to record time-use, our findings suggest that taking multiple observations—3 or 4 across the year—may be necessary to characterize a teacher's time-use practices in a broader sense. In addition, while it is clear that the typical half-day may look quite different from the typical full-day, focusing solely on average time-use may mask the underlying discretion that some teachers have to deviate from typical time-use. From a policy perspective, these uniquely complex features of time allocation in ECE settings should inform, and complicate, efforts to identify how time *should* be used to maximize students' readiness for K.

Finally, we zoom back out to the larger FDPK Study's goal of evaluating whether full-day pre-K is a promising policy lever for improving child outcomes. Previously published work on the FDPK Study documents short-term positive impacts of full- vs. half-day pre-K on a wide range of end-of-pre-K outcomes. If these promising findings persist as Study children move through school, policy makers will want to know how WPS teachers used that additional class time to the apparent benefit of students. This will be important because there are many different ways one could choose to allocate those extra hours, and it is not clear from prior research what the best choice would be. While we cannot definitively conclude that the observed changes to time-use in WPS caused the improvements in children's pre-K outcomes, the evidence from the current paper will make the FDPK findings more actionable for policy makers and may help generate hypotheses about how structural features of ECE shape children's learning opportunities.

Limitations and Lessons Learned

We conclude by acknowledging the limitations of the current study and opportunities for future work. First and foremost, this is a relatively small set of classes in a district that serves an important, but not nationally-representative, population of students. We cannot infer that teachers' time allocation in WPS represents what we would expect to find in ECE settings across the United States. Instead, our more modest goal is to illustrate our approach to studying time-use in ECE settings so that future researchers can be aware of the unique challenges they are likely to face in classrooms with pre-K aged children.

Second and relatedly, we were unable to design the FDPK Study with sufficient power to precisely estimate class-level associations between time-use and other outcomes. More specifically, we cannot conduct a mediation analysis to test whether any apparent causal impacts of full-day pre-K on child outcomes results from changes in time-use. In addition, the FDPK Study

team did not randomly assign teachers to full- vs. half-day classes. Instead, we asked WPS to follow its business-as-usual approach to teacher assignments in order to more closely reflect real-world district practices. Because our sample of teachers is small and not randomly assigned to full- vs. half-day classes, we cannot address the possibility that some of the differences in time-use could also be driven by differences in the background characteristics of those teachers. Focusing on teacher- or class-level variation in time-use seems like an obvious next step for ECE time-use researchers. We acknowledge, however, that exploring questions at the class level will require a study conducted at a much larger scale.

In terms of lessons learned, it is worth returning to the fact that the survey and observational instruments were not designed specifically to make cross-instrument comparisons. This was less of an issue for some activities that were simple to crosswalk (e.g., transitions, eating, and napping). However, we found that comparing unstructured play and structured play across observed and reported time use was less straightforward because reported time for content areas and activities were not mutually exclusive and did not allow for mixed content. In terms of lessons learned for future pre-K time-use surveys, we would suggest that researchers develop survey items that allow teachers to convey how/when they mix content together.

References

Author(s). (n.d.).

Bassok, D., & Latham, S. (2014). *Kids today: Changes in school-readiness in an early childhood era*. Annual Meeting of the Association for Public Policy Analysis and Management, Albuquerque, NM.

Bassok, D., Latham, S., & Rorem, A. (2016). Is Kindergarten the new first grade? *AERA Open*, *2*(1), 1–31. https://doi.org/10.1177/2332858415616358

Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: Who's minding the kids? In *Handbook of the Economics of Education* (Vol. 2, pp. 1163–1278). Elsevier. https://doi.org/10.1016/S1574-0692(06)02020-4

Burstein, L. (1995). *Validating national curriculum indicators*. Rand Corporation.

Camburn, E., & Barnes, C. A. (2004). Assessing the validity of a language arts instruction log through triangulation. *The Elementary School Journal*, *105*(1), 49–73. https://doi.org/10.1086/428802

Camburn, E., Correnti, R., & Taylor, J. (2001). *Examining differences in teachers' and researchers' understanding of an instructional log*. Annual Meeting of the American Education Research Association, Seattle, WA.

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., Clifford, R. M., Early, D. M., & Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in prekindergarten: Classroom engagement and school readiness gains. *Child Development*, *81*(5), 1534–1549. https://doi.org/10.1111/j.1467-8624.2010.01490.x

Coelho, V., Cadima, J., & Pinto, A. I. (2019). Child engagement in inclusive preschools: Contributions of classroom quality and activity setting. *Early Education and Development, 30*(6), 800–816. https://doi.org/10.1080/10409289.2019.1591046

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, *98*(4), 665–689. https://doi.org/10.1037/0022-0663.98.4.665

Currie, J., & Thomas, D. (1999). Does Head Start help Hispanic children? *Journal of Public Economics*, *74*(2), 235–262.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111–134.

Early, D. M., Barbarin, O., Bryant, D., Burchinal, M., Chang, F., Clifford, R., Crawford, G., Weaver, W., Howes, C., & Ritchie, S. (2005). Pre-Kindergarten in eleven states: NCEDL's multi-state study of pre-kindergarten and study of state-wide early education programs (SWEEP). *Informe Preliminar. Chapel Hill, NC: National Center for Early Development and Learning (NCEDL), Frank Porter Graham Child Development Institute*.

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D.-M. C., Crawford, G. M., Frome, P. M., Clifford, R. M., Burchinal, M., Howes, C., Bryant, D. M., & Pianta, R. C. (2010). How do pre-Kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *25*(2), 177–193. https://doi.org/10.1016/j.ecresq.2009.10.003

Elias, P., & Wheeler, P. (1976). Instructional activities as reported by teachers. *Journal of Teacher Education*, *27*(4), 326–328. https://doi.org/10.1177/002248717602700422

Engel, M., Claessens, A., Watts, T., & Farkas, G. (2016). Mathematics content coverage and student learning in Kindergarten. *Educational Researcher*, *45*(5), 293–300. https://doi.org/10.3102/0013189X16656841

Engel, M., Jacob, R., Claessens, A., & Erickson, A. (2021). Kindergarten in a large urban district. *Educational Researcher*, *50*(6), 401–415. https://doi.org/10.3102/0013189X211041586

Farran, D. C., Meador, D. M., Keene, A., Bilbrey, C., & Vorhaus, E. (2015). *Advanced narrative record manual* (2015 Edition). Vanderbilt University, Peabody Research Institute.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-Kindergarten on children's academic achievement. *The B.E. Journal of Economic Analysis & Policy*, *8*(1). https://doi.org/10.2202/1935-1682.1897

Fuligni, A. S., Howes, C., Huang, Y., Hong, S. S., & Lara-Cinisomo, S. (2012). Activity settings and daily routines in preschool classrooms: Diverse experiences in early learning settings for low-income children. *Early Childhood Research Quarterly*, *27*(2), 198–209. https://doi.org/10.1016/j.ecresq.2011.10.001

Golinkoff, D. G., Hirsh-Pasek, K., & Singer, R. M. (2006). *Play= Learning: How play motivates and enhances children's cognitive and social-emotional growth*. Oxford University Press.

Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. *Journal of Human Resources*, *40*(3), 533–558. https://doi.org/10.3368/jhr.XL.3.533

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

La Paro, K. M., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., Bryant, D., Early, D., Clifford, R., Barbarin, O., Howes, C., & Burchinal, M. (2009). Quality in Kindergarten classrooms:

Observational evidence for the need to increase children's learning opportunities in early education classrooms. *Early Education & Development*, *20*(4), 657–692. https://doi.org/10.1080/10409280802541965

Li, W., Duncan, G. J., Magnuson, K., Schindler, H. S., Yoshikawa, H., & Leak, J. (2020). *Timing in early childhood education: How cognitive and achievement program impacts vary by starting age, program duration, and time since the end of the program* [EdWorkingPaper: 20-201]. EdWorkingPapers.com. https://www.edworkingpapers.com/ai20-201

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*, 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, *26*(1), 33–51.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, *21*(1), 29–45.

National Association for the Education of Young Children. (2020). *Developmentally Appropriate Practice: A Position Statement of the National Association for the Education of Young Children* (pp. 1–44).

Nores, M., Belfield, C. R., Barnett, W. S., & Schweinhart, L. (2005). Updating the economic impacts of the High/Scope Perry preschool program. *Educational Evaluation and Policy Analysis*, *27*(3), 245–261. https://doi.org/10.3102/01623737027003245

Phillips, D. A., Gormley, W. T., & Lowenstein, A. E. (2009). Inside the pre-Kindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. *Early*

Childhood Research Quarterly, *24*(3), 213–228. https://doi.org/10.1016/j.ecresq.2009.05.002

Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., Duncan, G. J., Dynarski, M., Magnuson, K. A., & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-Kindergarten effects a consensus statement*. Brookings Institution. https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf

Pianta, R., 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

Piasta, S. B., Pelatti, C. Y., & Miller, H. L. (2014). Mathematics and science learning opportunities in preschool classrooms. *Early Education and Development*, *25*(4), 445–468. https://doi.org/10.1080/10409289.2013.817753

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study: Final report* (No. 2012–45; OPRE). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Jang, H., & Chavez, B. (2021). *Stanford education data archive (version 3.0)*. http://purl.stanford.edu/db586ns4974

Ritchie, S., Howes, C., Kraft-Sayre, M., & Weiser, B. (2001). Emerging academics snapshot. *Unpublished Measure, University of California at Los Angeles*.

Rowan, B., & Correnti, R. (2009). Studying reading instruction with teacher logs: Lessons from the Study of Instructional Improvement. *Educational Researcher, 38*(2), 120–131. https://doi.org/10.3102/0013189X09332375

Shonkoff, J. P., Phillips, D. A., & National Research Council (U.S.) (Eds.). (2000). *From neurons to neighborhoods: The science of early child development*. National Academy Press.

Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the Enacted Curriculum, the Reform Up Close Study* (Issue 31). Consortium for Policy Research in Education.

Stipek, D., Feller, R., Daniels, D., & Milburn, S. (1995). Effects of different instructional approaches on young children's achievement and motivation. *Child Development, 66*(1), 209–223.

U.S. Department of Education, National Center for Education Statistics. (2016). *Common core of data*. http://nces.ed.gov/ccd/elsi/

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84*(6), 2112–2130. https://doi.org/10.1111/cdev.12099

Westminster Public Schools. (2021). *Early Childhood Education / About ECE*. https://www.westminsterpublicschools.org/Page/8793

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-Kindergarten programs: Evaluation of five state pre-Kindergarten programs. *Journal of Policy Analysis and Management, 27*(1), 122–154. https://doi.org/10.1002/pam.20310

Tables

Table 1. Summary of Key Literature: Lessons in Accurately Measuring Classroom Time Use

				Instrument Type(s)				Agreement Rate(s)		
Author	**Year**	**Dataset**	**Grade Level**	*Obs.*	*Surveys*	*Logs*	*Other*	*Math*	*ELA*	*Activities*
Current Study	*2021*	*Pre-K Teachers in Colorado School District*	*Pre-K: Half-Day*	*X*	*X*			0.17[a]	-0.33[a]	--
			Pre-K: Full-Day	*X*	*X*			0.49[a]	-0.04[a]	--
Elias & Wheeler	1976	Beginning Teacher Evaluation Study (BTES)	Elem.	X		X		0.47[a]	0.57[a]	--
Smithson & Porter	1994	Reform Up Close (RUC) Study	HS	X		X		47-80%	--	61-83%
					X	X		-0.05-0.93[a]	--	0.25-0.63[a]
Burstein	1995	Math teachers from California and Washington	MS & HS		X	X		83%	--	35-60%
					X		X	42-96%	--	--
Mayer	1999	Anonymous district	HS	X	X			0.85[a]	--	--
Camburn & Barnes	2004	Pilot study of language arts log by CPRE	Elem.	X		X		--	73%	--
Rowan & Correnti	2009	Study of Instructional Improvement (SII)	Elem.		X	X		--	--	--

FN: The superscript [a] indicates agreement statistics that were reported as correlations, and % sign indicates agreement statistics that were reported as percent agreement rates. Elias & Wheeler (1976) Math and ELA = p. 328; Smithson & Porter (1994) Math = Row1: Table 1, p. 10 and Row2: Table 3, p. 12; Smithson & Porter (1994) Instructional Activities = Table 2, p. 11; Burstein (1995) Math = Row1: p. 32, Row2: Table 3.2, p. 29 and Table 3.3, p. 30 ; Burstein (1995) Instructional Activities = Table 5.3, p. 53; Mayer (1999) = p.43; Camburn & Barnes (2004) = p. 56

Table 2. Overview of Class Observation and Survey Data Collection, with Counts of Number Collected

		Counts		Obs Windows in Pre-K Year					Survey
		Classrooms	Classes	Oct - Dec	Jan - Feb	Feb - Mar	Apr - May	Total	Apr - May
Cohort 2	Full-Day	9	9	17 of 17		15 of 17	15 of 17	*47 of 51*	*12 of 13*
	Half-Day	4	8						
	Cohort Total	13	17						
Cohort 3	Full-Day	9	9	16 of 17	17 of 17	17 of 17	17 of 17	*67 of 68*	*12 of 13*
	Half-Day	4	8						
	Cohort Total	13	17						
Cohorts 2 + 3	Full-Day	18	18	33 of 34	17 of 17	32 of 34	32 of 34	*114 of 119*	*24 of 26*
	Half-Day	8	16						
	Overall Total	26	34						

FN: Cell contents report the number of data points collected out of the total we intended to collect based on the study design. The first column, labeled 'Classrooms', counts the number of unique teachers' classrooms by cohort. This is slightly different than the count of 'Classes' that appears in the second column, which counts the number of unique sets of students for a given teacher-year. For example, in a given year/cohort, a half-day teacher has one classroom, but two classes of students (AM and PM). 'Obs' stands for 'observation' and indicates that the study design called for observing each classroom once during the window (covering both the complete AM and PM session for half-day classrooms). 'Survey' indicates that the study design called for asking the teacher in each classroom to report their typical time use on an end-of-year survey. For instance, there were 34 total classes (across half- and full-day in both cohorts). We completed 114 of the 119 observations we set out to capture, and we collected 24 of the 26 surveys we sent to classroom teachers (two full-day teachers failed to return the survey).

Table 3. Lead Teacher Characteristics: Teachers Assigned to Half- and Full-Day Pre-K Classes

Teacher Characteristic	Statistic	Overall	Half-Day	Full-Day	Difference P-Value
Female	%	100%	100%	100%	0.000
	N	16	6	10	
Asian	%	6%	17%	0%	0.322
	N	1	1	0	
Hispanic	%	13%	17%	10%	0.733
	N	2	1	1	
White	%	75%	50%	90%	0.117
	N	12	3	9	
Years of Experience	Mean	15.9	18.3	14.4	0.459
Has a BA+	%	88%	67%	100%	0.126
	N	14	4	10	
Has No Certifications	%	0%	0%	0%	0.000
	N	0	0	0	
Total Number of Teachers		**16**	**6**	**10**	

FN: Table 3 compares the teacher characteristics of the half-day classes (AM and PM) in Cohorts 2 and 3 to the full-day classes in Cohorts 2 and 3 by reporting the number of lead teachers in each category (except for the category "mean # of years in public pre-K"). These data were collected via the teacher survey administered at the end of each school year in each cohort and include the total number of unique teachers. See Table 4 for more details about these teachers' placements.

Table 4. Frequencies of classroom observations and survey collection by cohort, teacher, and class type.

Teacher (N=16)	# of Classrooms per Cohort (N=22)	Data Source	Cohort 2	Cohort 3	Teacher Totals
A	2: AM/PM	Survey	1	1	2
		# Times Observed	5	8	13
B	2: AM/PM	Survey	1	1	2
		# Times Observed	6	8	14
C	2: AM/PM	Survey	--	1	1
		# Times Observed	--	8	8
D	2: AM/PM	Survey	1	--	1
		# Times Observed	6	--	6
E	2: AM/PM	Survey	1	--	1
		# Times Observed	4	--	4
F	2: AM/PM	Survey	--	1	1
		# Times Observed	--	8	8
G	1: Full	Survey	1	1	2
		# Times Observed	3	4	7
H	1: Full	Survey	1	1	2
		# Times Observed	3	4	7
I	1: Full	Survey	1	1	2
		# Times Observed	2	4	6
J	1: Full	Survey	1	1	2
		# Times Observed	3	4	7
K	1: Full	Survey	1	1	2
		# Times Observed	3	4	7
L	1: Full	Survey	1	--	1
		# Times Observed	3	--	3
M	1: Full	Survey	1	Missing	1
		# Times Observed	3	4	7
N	1: Full	Survey	1	1	2
		# Times Observed	3	4	7
O	1: Full	Survey	Missing	1	1
		# Times Observed	3	4	7
P	1: Full	Survey	--	1	1
		# Times Observed	--	3	3
Cohort Totals		**Survey**	**12**	**12**	**24**
		# Times Observed	**47**	**67**	**114**
		Total # Planned Obs.	51	68	119
		% Planned Obs. Completed	*92.2*	*98.5*	*95.8*

FN: An "--" indicates that the given teacher was not present in WPS in the given cohort year; "Missing" indicates a missing teacher survey; The max number of entire class-period observations logged in Cohort 2 was three: (1) Nov/Dec, (2) Feb/Mar, and (3) Apr/May. The max number of entire class-period observations logged in Cohort 3 was four: (1) Oct/Nov, (2) Jan/Feb, (3) Feb/Mar, and (4) Apr/May. Classroom observations were not collected for Cohort 1. Observational rubrics were applied by in-person observers in Cohort 2 and were applied to classroom video recordings in Cohort 3.

Table 5. Crosswalk between Time-Use Codes from the Survey and the Activity/Content Codes from the Observational Protocol.

Teacher Survey	Classroom Observation Protocol
Response Option	**Response Option**
Reading and Language Arts (ELA)	Reading Language Arts Literacy
Math	Math
Social Studies or Science	Science Social Studies
Art/Music/Drama/Dance	Music Visual Arts Specials
Structured play (teacher-directed)	Gross Motor Small Group + Centers Small Groups Only
Unstructured play (child-directed, including recess)	Playground Centers only
Transitioning between Activities	Transition
Eating	Meal and snacks
Napping or Quiet Resting	Napping
Other	Other content
* N/A - These activities overlap with the other crosswalked categories and are left out of the RQ3 analysis.	No Content Observed Content Not Applicable Whole group

Table 6. Difference in the Number of Minutes Spent Per Day on a Given Subject, between Obervational Data and Teacher Self-Reported Survey, Separately by Class Type (Half- vs. Full)

	Half-Day $n=106$		Full-Day $n=108$	
Subject	*Obs. Mean*	$\widehat{\beta_1}$	*Obs. Mean*	$\widehat{\beta_1}$
ELA	30.7	-11.6 *	50.6	-8.9
Math	12.3	7.0	19.3	7.7
Science & Soc. Studies	28.4	-16.9 **	41.7	-18.8 *
Music & Art	25.9	-13.9 ***	39.0	-17.4 *
Structured Play	35.3	-12.5 *	45.1	-11.8 *
Unstructured Play	48.6	7.3	88.3	12.2
Transitions	37.0	-22.1 ***	71.0	-44.6 ***
Eating	13.2	5.6 **	26.5	24.2 ***
Napping	0.0	0.0	69.7	-5.7

*FN: * < = 0.05, **<=0.01, *** <=0.001. "Obs. Mean" presents the average number of minutes spent on the given subject according to the* observational *rubrics (e.g., on average half-day teachers were observed spending 30.7 minutes per day on ELA). β_1 represents how many more/fewer minutes were spent on the subject according to the survey for the same teacher's class in the same year. For instance, on average half-day teachers reported on their* surveys *spending 11.6 fewer minutes on ELA than was observed on their corresponding observations from the same year. The estimates of β_1 come from 18 separate class-by-year fixed effect regression models (nine subjects, separately for half- and full-day classes) with the following specification: The outcome is the number of minutes spent on a given subject in observation window w, in class c, in year y, according to source s. β_1 is the coefficient on an indicator variable, coded 0 if time spent was reported from an observational rubric, and coded 1 if time spent was reported from an annual teacher survey. The Ns reflect two sources of data multiplied by the 53 and 54 observations across class-years (half- and full-day, respectively) that could be matched to annual survey data collected at the classroom level.*

Table 7. Descriptive statistics in average percent per day from observed time use collected during classroom observations

	Half Day (AM & PM) n = 16				Full Day n = 18			
Activity Type	Mean	SD	p25	p75	Mean	SD	p25	p75
Napping	0.0	0.0	0	0	19.3	3.4	17	22
Eating	8.5	2.6	7	9	8.1	4.2	4	11
Transitions	19.5	4.5	16	23	20.1	4.3	18	23
Playground	12.7	4.5	10	16	10.4	3.0	8	12
Whole Group	26.7	6.7	22	31	16.8	3.7	15	19
Small Groups	5.4	6.5	1	8	6.5	4.0	3	9
Small Group & Centers	10.3	7.0	6	16	2.9	3.8	0	4
Centers	13.7	6.9	9	18	13.7	5.6	10	16
Specials	3.2	3.5	0	4	2.2	2.7	0	3
	100.0				*100.0*			
Content	Mean	SD	p25	p75	Mean	SD	p25	p75
Not applicable	23.2	14.8	7	35	37.4	10.2	34	42
No Content Observed	6.9	10.2	0	11	9.2	10.0	0	19
Mixed Content	47.8	10.9	39	52	34.7	8.4	29	41
ELA	5.8	5.1	1	10	4.4	4.5	2	5
Science & Soc. Stud.	4.0	2.9	2	5	4.2	3.3	1	6
Music and Visual Arts	2.7	2.2	0	5	4.1	3.7	1	5
Gross Motor	2.7	4.0	0	5	1.7	2.4	0	4
Math	1.6	2.3	0	3	0.6	0.8	0	1
Other Content	5.4	4.4	1	8	3.6	2.9	1	6
	100.0				*100.0*			

FN: Cell contents report the percent of time per day observed on each activity type (upper panel) or content area (lower panel). Results for half-day classes are shown on the left (N=16 total class-years), and full-day classes are shown on the right (N=18). Each panel totals 100% of the observed school day.

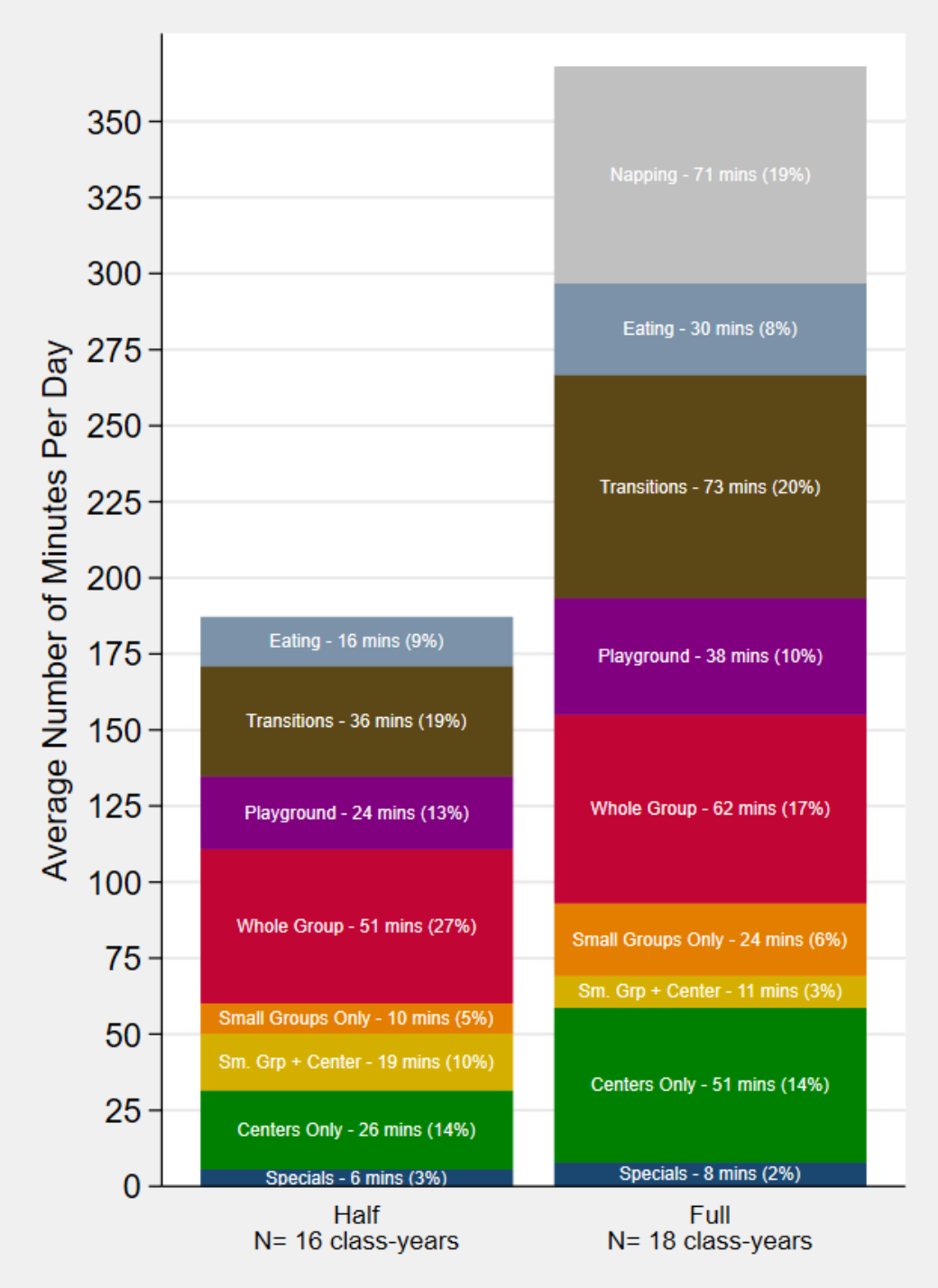
Table 8. Descriptive statistics for ten most frequently observed content combinations of mixed content across observation windows, classrooms, and cohorts

	Half Day (AM & PM)					Full Day				
	Number of Times	Descriptives for Minutes Spent per Day in a Given Content Combination				Number of Times	Descriptives for Minutes Spent per Day in a Given Content Combination			
Content Combinations within Mixed Content	Combination Observed	Average Minutes	SD of Minutes	Minimum Minutes	Maximum Minutes	Combination Observed	Average Minutes	SD of Minutes	Minimum Minutes	Maximum Minutes
ELA + Soc. Stu./Science	65	18.7	7.3	5	38	43	15.2	6.2	5	38
ELA + Music/Art	51	19.5	10.4	5	49	42	13.7	3.2	9	20
Soc. Stu./Science + Music/Art	50	14.2	6.2	5	25	32	15.8	9.1	6	41
ELA + Math	44	18.1	7.0	3	35	24	12.7	6.2	3	23
Music/Art + Other	36	25.9	10.5	5	58	25	19.3	6.6	8	40
ELA + Other	28	18.6	12.1	5	48	17	18.5	10.9	2	45
Soc. Stu./Science + Other	19	15.1	5.8	10	28	17	16.8	13.4	5	55
Math + Music/Art	18	28.2	19.9	2	75	21	18.6	6.8	8	35
ELA + Gross Motor	17	13.4	5.0	4	19	13	15.5	7.7	3	27
Math + Other	11	24.4	18.2	6	69	8	22.5	15.0	4	44

Table is limited to time-use episodes that are coded as more than one simultaneous focal content area. The table reports descriptive statistics (count, mean, standard deviation ("SD"), minimum, and maximum) of distinct time-use episodes that are coded as one of the top ten most common content combinations. 92% of all mixed content episodes are contained in this top ten. Rows are sorted from most- to least- frequent combinations among half-day classes.

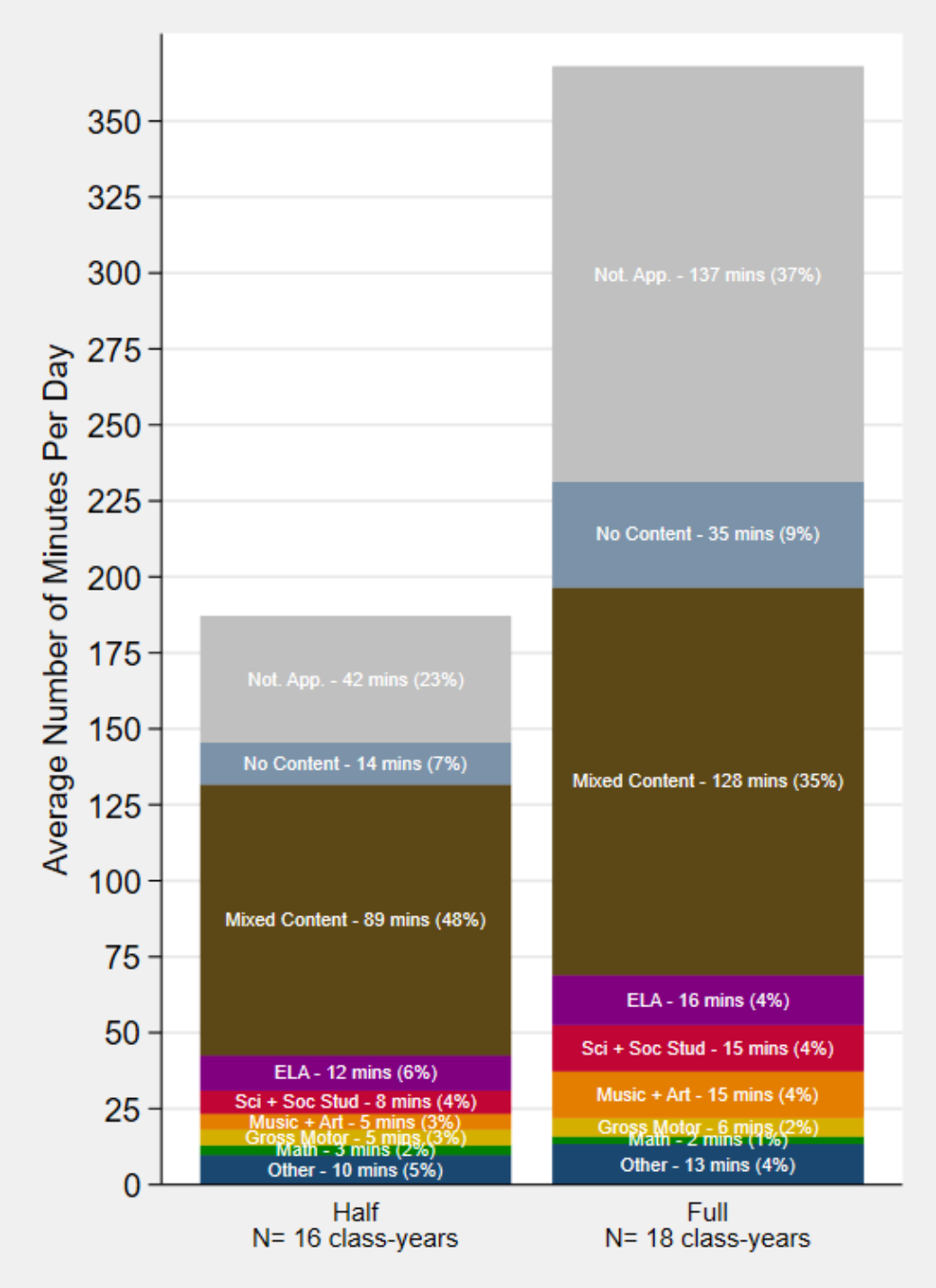
Figures

Figure 1. Average Minutes Observed per Day in Mutually-Exclusive, Activity Types, by Half- and Full-Day



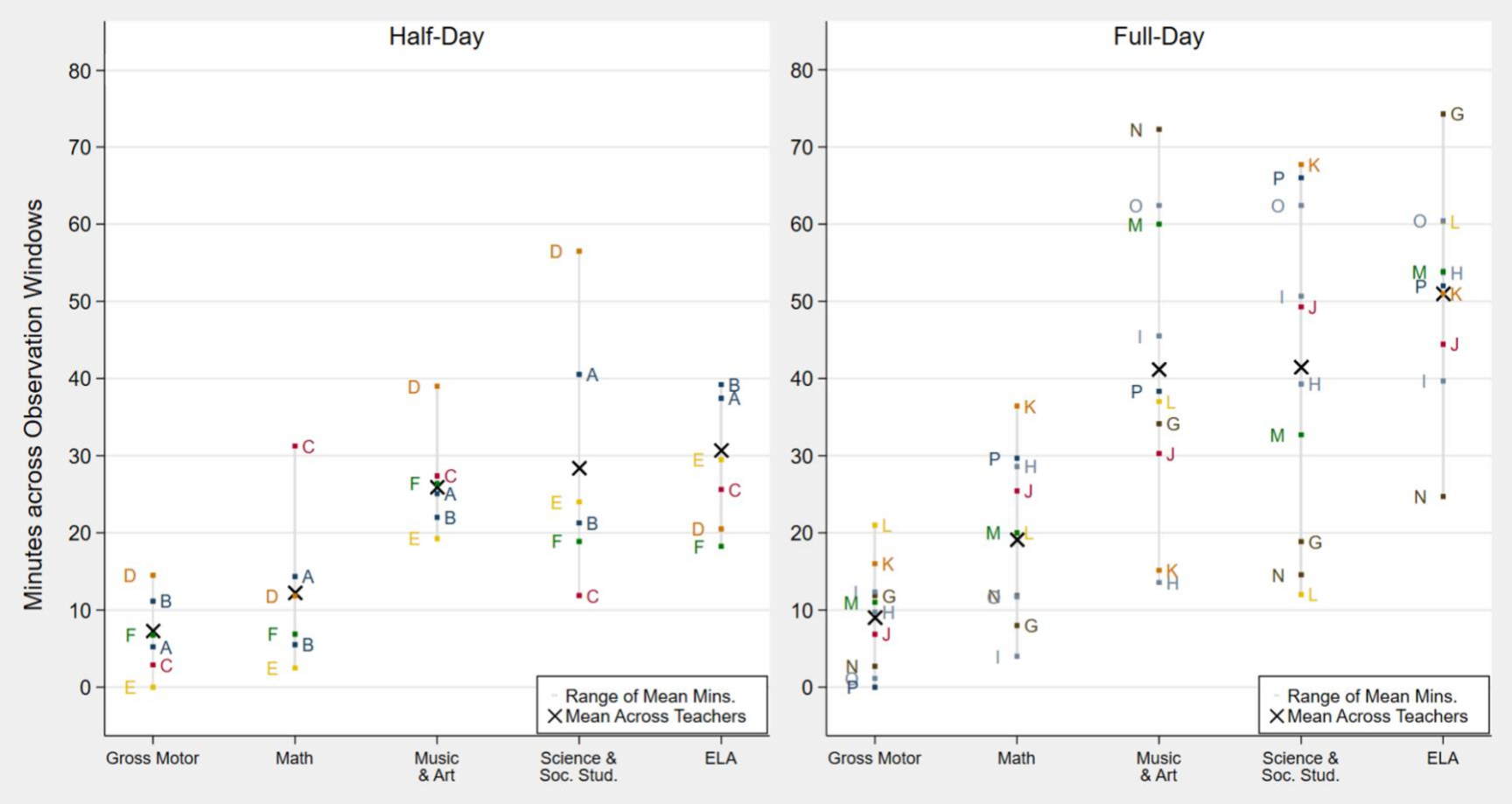
FN: Only data collected from classroom observations were used in the stacked bar graphs. Data are pooled across cohorts, by presented separately by classtype. This results in 16 unique half-day class-years and 18 unique full-day class-years, where class-years are identical to the 'classes' described in Table 2.

Figure 2. Average Minutes Observed per Day on Mutually Exclusive Content Areas, by Half- and Full-Day



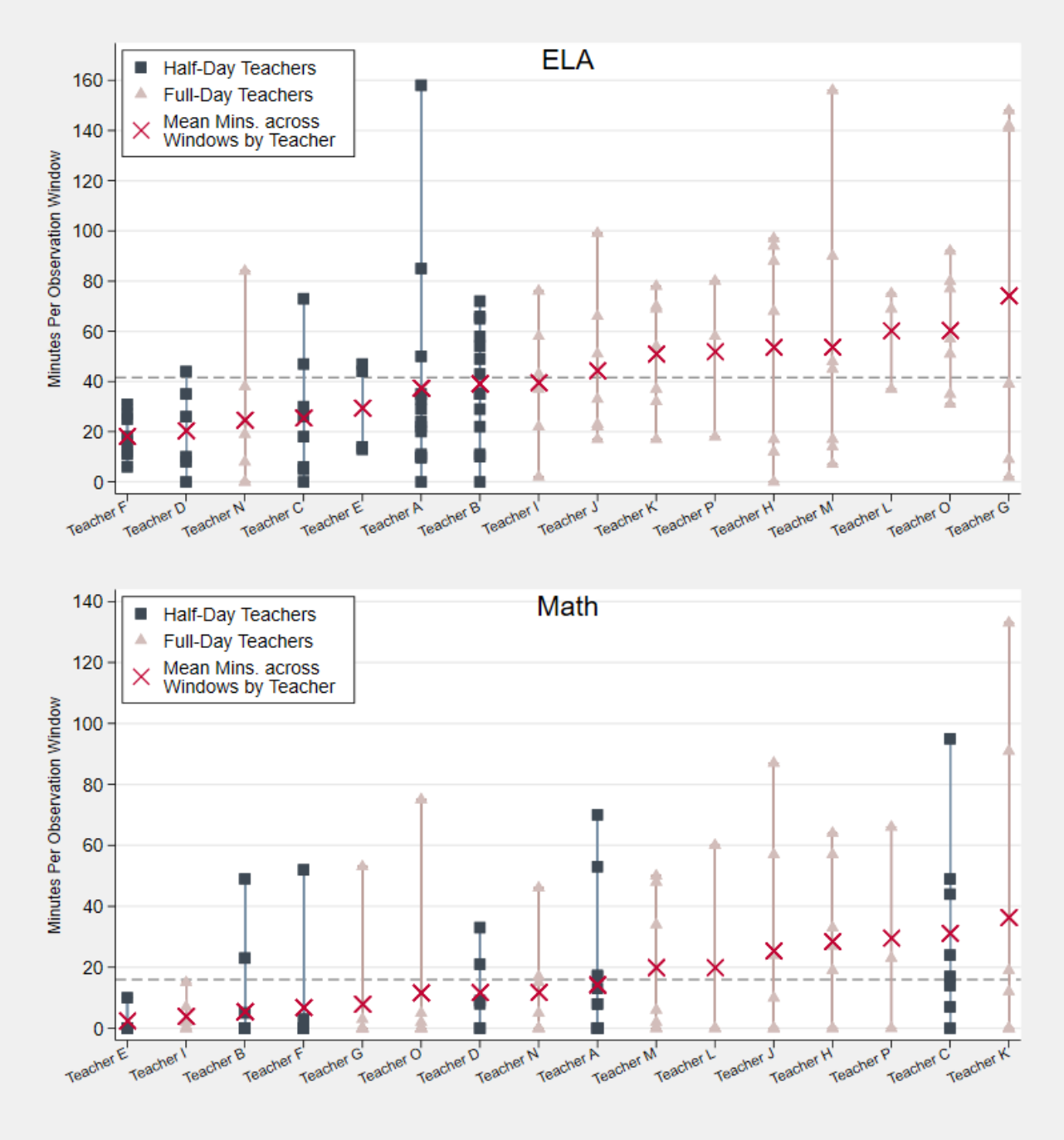
FN: Only data collected from classroom observations were used in the stacked bar graphs. Data are pooled across cohorts, by presented separately by classtype. This results in 16 unique half-day class-years and 18 unique full-day class-years, where class-years are identical to the 'classes' described in Table 2.

Figure 3. Variation across Teachers in Minutes Spent on Five Content Areas, Separately for Half- and Full-Day Classes



FN: Each colored dot represents a different teacher and has a letter identifier that corresponds to the Teacher letters in Table 2. The Y-axis captures the average number of minutes the given teacher was observed spending on the given content area (averaged across all her observations). The grey vertical line illustrates the range of minutes spent on a given content area across teachers. The larger, black "×" on each content area's vertical line represents the mean minutes spent on that subject overall.

Figure 4. ELA & Math: Variation across Each Teacher's Observation Windows in Minutes Spent on Given Content Area, Separately for Half- and Full-Day Classes



FN: Teachers (x-axis) were observed for an average of 6 complete class periods over the two-year time frame (min=3, max=14). Each dot represents an entire class-period observation. The red "×" on each teacher's vertical line represents the mean minutes that teacher spent across all her observations on the given content area. The dashed grey line indicates mean time spent on given content area overall (across all teachers and observation windows).

Online Appendix A: School-Level Descriptive Statistics for Students

Table A9. Demographic Characteristics of Pre-K Students Across 10 Study Sites

	# of Class-Years		Gender	Race/Ethnicity			%	% Home	Average
			%	%	%	%	FRPL	Language	Child
School Site	Half	Full	Male	Asian	Hispanic	White	Eligible	English	Age
School Site 1	12	0	55%	4%	72%	20%	67%	49%	4.51
School Site 2	4	2	53%	4%	66%	22%	74%	51%	4.48
School Site 3	0	2	52%	0%	62%	31%	55%	55%	4.49
School Site 4	0	2	50%	0%	81%	16%	75%	44%	4.56
School Site 5	0	2	69%	0%	72%	25%	66%	59%	4.56
School Site 6	0	2	47%	6%	53%	41%	59%	59%	4.47
School Site 7	0	2	56%	0%	69%	25%	63%	50%	4.53
School Site 8	0	2	48%	0%	74%	19%	65%	61%	4.56
School Site 9	0	2	44%	3%	75%	16%	75%	44%	4.49
School Site 10	0	2	53%	6%	59%	25%	72%	75%	4.53
Overall	16	18	54%	3%	70%	22%	68%	51%	4.52

Note. N = 34 classes of students in current study (Cohorts 2 and 3).

Online Appendix B: Classroom Observation Protocol Data Details

Members of our team collected data about pre-K time-use via classroom observations in the second and third cohorts. In both cohorts, team members conducted classroom observations using a time-use protocol based on the *Advanced Narrative Record Observation for Classrooms* developed by Farran and colleagues (2015). Observers recorded a beginning and an ending time for each time-use episode in addition to the following: the type of activity observed (see online Appendix C for list of activity types), if instruction was led by the lead teacher, the number of small groups if applicable, level of instruction (no instruction, low level of instruction, basic skill instruction, some inferential learning, or high inferential learning); students' engagement level (no instruction, low, medium low, medium, medium high, high); and students' compliance level (low, medium low, medium, medium high, and high), as well as any notes observers took for the specific time-use episode. Observers also filled in the observation date, the average number of students present during the observation, the average number of teachers present during the observation, the average number of teaching assistants present during the observation, and the average number of parents or volunteers present during the observation. See the example presented in Figure B1.

Teacher Name:	School Name:	AM, PM, Full	Observation Date:	Avg. # Children	Avg. # Teachers	Avg. # Assistants	Avg. # Parents/ Vols	Observer Name:
Teacher, A (Rm 1)	Hodgkins	Half-PM	4/22/18	11	1	1	0	Observer #1

					If Instruction Occurring...			Overall Classroom Observations			
Start Time	**End Time**	**Activity Type**	**Is Instruction Teacher Led?**	**Number of Small Groups**	**Content Area 1**	**Content Area 2**	**Content Area 3**	**Level of Instruction**	**Students' Engagement Level**	**Students' Compliance Level**	**Notes and/or Description**
12:00 PM	12:22 PM	Whole Group	Yes	~~"n/a"~~	Science	Art	None	2-Basic Skill Instruction	5-High	4-Medium High	As children enter the classroom, they can choose if they want to play a game that matches adult animals to their babies, or draw pictures of animal homes, or play with bamboo blocks at tables.
12:22 PM	12:25 PM	Transition	No	~~"n/a"~~	None	None	None	~~N/A - No Instruction~~	3-Medium	4-Medium High	Children clean up the toys and find aspot on the floor where they can sit.
12:25 PM	12:39 PM	Whole Group	Yes	2	Science	Literacy	<<choose one if applicable>>	3-Some Inferential Learning	5-High	5-High	Half the class leaves to go to the bathroom. The other group talks in at the carpet about the animals they saw at th zoo during their fieldtrip the previous day. Then they write 'yes' or 'no' by each animal based on whether or not they saw the animal.
12:39 PM	12:53 PM	Whole Group	Yes	~~n/a~~	Science	Art	<<choose one if applicable>>	3-Some Inferential Learning	4-Medium High	5-High	Then, another group leaves to go to the bathroom. The returning group works on a worksheet with the animals where they color the animals who are on the worksheet they saw at the zoo. Teacher asks questions about the trip yesterday. Did you see an elephant?What did you notice about the polar bear?

Figure B1. Example of completed narrative record using an adapted form of the *Advanced Narrative Record* protocol (Farran et al., 2015).

We created a set of analytic variables that documented the average time a teacher spent on each primary content area and each activity in a given class-year across observation windows. First, we summed elapsed time from each time-use episode within a cohort, class and observation window by content area and activity type. This set of elapsed time by content/activity observed *within* an observation window was then averaged across observation windows, classrooms, and cohorts to create an average of the total daily time teachers in our study devoted to each activity or content area. Our research team recorded a primary content area, in addition to an optional secondary content area and tertiary content area, that may have been observed in each time-use episode. In other words, every episode could have from zero to three content areas observed, where zero content areas would correspond to "no instruction" happening (i.e., during transitions). Each episode always had one activity recorded. Total time spent on content areas within an observation window was calculated separately for instances in which the content area appeared in each of the primary, secondary, and tertiary content categories *and* for instances where the content appeared in *any* of the three subject categories. Neither the pooled content area totals nor the secondary/tertiary content categories are used in the present analysis (except to indicate that an episode should be classified as *mixed*). Across the two cohorts, we are missing 16 time-use episodes because observers did not include an end time for the episode in the narrative record.

To create the set of variables reporting the proportion of a day devoted to each content area and activity, we used the average total time for an observation for half- and full-day pre-K as the denominators. The numerator of this proportion, again calculated by class-type, was the average observed time for each content type or activity type across observation windows, classrooms, and cohorts. The resulting variable is the average proportion of time per day devoted to a particular content area or activity by class-type. Since we had at least two observations for each class-year

and we focused on the "average" time across all observation windows, we did not impute for missing data. Tables B1 and B2 present descriptive statistics for classroom observations by cohort, class-type, and subject in mean minutes per day.

Table B1. Descriptive statistics for Cohort 2 by class-type and subject from classroom observations in mean minutes per day

	Half Day (AM & PM) $n = 8$				Full Day $n = 9$			
Activity Type	Mean	SD	p25	p75	Mean	SD	p25	p75
Napping	0.0	0.0	0.0	0.0	75.3	12.2	65.3	88.8
Eating	19.6	8.2	15.0	22.0	42.6	14.9	37.7	58.6
Transitions	33.5	8.8	24.8	39.1	66.1	13.7	60.6	75.1
Playground	21.9	13.0	12.3	28.5	41.3	8.9	35.7	46.2
Whole Group	55.7	24.3	38.5	61.9	65.1	19.6	57.4	73.5
Small Groups	8.0	15.5	0.0	7.0	20.3	12.7	13.3	26.5
Small Group & Centers	15.0	15.5	0.0	23.7	12.1	18.3	0.6	14.7
Centers	36.3	11.1	27.4	47.2	52.1	26.2	31.6	59.1
Specials	4.2	7.0	0.0	6.7	5.4	4.3	3.7	7.0
Content Areas	Mean	SD	p25	p75	Mean	SD	p25	p75
Not applicable	21.1	20.1	5.2	35.6	119.1	33.5	92.5	138.5
No Content Observed	24.9	25.2	2.1	42.5	64.4	33.5	50.5	80.4
Mixed Content	96.9	25.1	77.3	113.8	116.2	35.7	96.5	128.1
ELA	16.6	15.6	4.5	21.9	25.3	20.7	17.0	21.3
Science & Soc. Stud.	8.5	5.6	5.0	11.7	7.1	6.9	1.4	13.8
Music and Visual Arts	7.6	4.7	4.5	11.8	19.5	17.4	13.3	17.8
Gross Motor	8.8	9.9	0.0	16.9	9.9	10.0	0.0	15.3
Math	2.9	5.3	0.0	2.7	4.4	3.7	1.1	6.0
Other Content	7.1	8.5	0.0	12.6	14.3	13.9	3.0	28.0

FN: Cell contents report the average number of minutes per day observed on each activity type (upper panel) or content area (lower panel). Results for half-day classes are shown on the left (N=8 total classes in Cohort 2), and full-day classes are shown on the right (N=9). Each panel totals 100% of the observed school day.

Table B2. Descriptive statistics for Cohort 3 by class-type and subject from classroom observations in mean minutes per day

	Half Day (AM & PM) $n = 8$				Full Day $n = 9$			
Activity Type	Mean	SD	p25	p75	Mean	SD	p25	p75
Napping	0.0	0.0	0.0	0.0	66.9	12.7	60.5	68.3
Eating	12.6	3.0	10.4	14.7	17.8	7.8	12.1	23.9
Transitions	38.8	5.1	33.9	43.1	80.6	13.0	73.5	82.2
Playground	26.0	5.8	21.2	29.7	35.4	11.6	26.6	43.5
Whole Group	45.6	9.1	39.9	52.7	58.9	10.1	53.1	61.0
Small Groups	12.1	8.6	5.3	21.6	27.3	16.6	11.8	36.4
Small Group & Centers	22.1	8.2	14.0	29.5	9.1	7.8	2.2	12.2
Centers	15.7	8.5	10.7	18.2	49.4	18.0	38.7	57.1
Specials	7.1	4.8	4.0	10.0	10.5	12.5	0.0	13.0
Content Areas	Mean	SD	p25	p75	Mean	SD	p25	p75
Not applicable	62.2	10.7	53.2	70.6	154.1	29.1	124.4	179.3
No Content Observed	2.8	4.5	0.0	4.1	5.2	7.2	0.6	8.2
Mixed Content	81.3	13.1	73.7	91.4	138.9	25.0	117.5	160.8
ELA	6.6	5.5	1.2	11.3	7.6	3.2	4.1	10.3
Science & Soc. Stud.	6.7	5.7	3.0	10.0	23.3	9.6	16.7	33.5
Music and Visual Arts	2.7	3.0	0.4	5.4	11.2	10.0	1.1	19.4
Gross Motor	1.7	3.3	0.0	2.5	2.3	5.3	0.0	0.0
Math	3.5	4.9	0.0	8.8	0.6	1.0	0.0	0.6
Other Content	12.6	7.1	7.5	17.6	12.7	8.3	5.7	21.3

FN: Cell contents report the average number of minutes per day observed on each activity type (upper panel) or content area (lower panel). Results for half-day classes are shown on the left (N=8 total classes in Cohort 3), and full-day classes are shown on the right (N=9). Each panel totals 100% of the observed school day.

Online Appendix C: Observation Protocol's Activity and Content Code Details

Activity Codes:

1) *Meals and Snacks.* Mealtime is coded when students are scheduled to eat breakfast, lunch, or snacks.
2) *Transitions.* Anytime the teacher and students are waiting for the next activity, going to the bathroom, lining up for lunch, etc.
3) *Napping.* Children are resting or sleeping.
4) *Playground.* Code playground when children are playing on the playground or in another outside area designated for outdoor play/recess.
5) *Whole Group*. When all children are working on the same activity, regardless of seating arrangement
6) *Small Group Only (No Centers).* Children are assigned to small groups (non are self-selected), and groups are working on different activities.
7) *Small Group + Centers.* Some children are assigned to small groups, and other children have self-selected a center. Groups are working on different activities.
8) *Centers Only (No Small Groups).* All children are working in self-selected groups in centers (or tabletop centers) with a choice of activities. All children must either choose their center or the activity/materials they engage with in order to code centers.
9) *Specials.* Some classes participate in special activities led by "Specials" teachers either inside or outside the classroom, such as music, art, gardening and/or computer. When children are participating in these activities, code activity type as SP. Use this code when an adult other than the classroom teacher or assistant is leading an activity.

Content Codes:

The classroom observation protocol included the following instructional content codes that were defined as follows:

- *Language Arts*. Learning letter names and sounds, tracing letters or numerals, writing with a focus on letter or symbol formation, rhyming, alphabet songs, etc. Focus is on the parts or components of written or spoken language.
- *Literacy*. (Often appears with Reading, Language Arts). Knowing the days of the week and months of the year (including songs in which the teacher or assistant point to written words for the days of the week or months of the year), writing name, conversations or discussions for the purpose of facilitating language development where there is no other primary focus.
- *Reading*. Connected text with meaning, reading a book or reading other text that does not involve a book (e.g., morning message from the board) with or without asking children questions.
- *Math*. Numbers, counting, comparing quantity, sorting (by size, weight, length, etc.), graphing, patterning, telling time, shape identification, positional words

(above, below), measurement, songs with math content (e.g., 5 Little Monkeys, songs about left and right).

- *Science*. Health, nature, food (unless focused on learning about a specific culture), physical properties, biology, earth and space, seasons, weather, using scientific tools (e.g. rain gauge), animals, life cycles, color identification, sorting by color, mixing colors, vehicles (when talking ut the parts and mechanics), songs with science content (e.g., Hokey Pokey, Head Shoulders Knees Toes, Old McDonald), etc.
- *Social Studies*. Community helpers, getting along with others, socio-emotional concepts, geography, history, vehicles (when talking about moving goods in a community context or the occupation of being a driver), songs with social studies content (e.g., Good Morning to You for learning to greet, Weather Watcher, What is the Weather Like Today, songs to learn names of children).
- *Music/Movement*. Music (e.g., Tootie Ta, If You're Happy and You Know It, Where is Thumbkin, Twinkle Twinkle Little Star, Bear Hunt, Wheels on the Bus, BINGO, Shake it Like a Duck), Good morning songs when there is no greeting/interaction component, movement games (e.g. Freeze game), playing instruments.
- *Art*. Art projects, painting, coloring, collage making, and drawing with crayons.
- *Gross Motor.* Activities relating to large muscle movement and control, such as bean bag toss, riding toys, balance beam, typically takes place outside or in the gym.
- *Other Content.* Activities such as blocks, toys, fine motor materials, non-educational TV programs with no clearly defined instructional content.
- *Not Applicable.* Not a code used by the observer. This code is used when the observer did not select any of the other content dropdown options in the protocol.
- *Missing.* Not a code used by the observer. This code indicates time-use episodes in which the observer left the content area blank (i.e. the observer did not choose "No Content Observed" nor did they leave the content area as the default "Not Applicable")
- *No Content Observed*. Directions with no educational content, most meal and nap episodes, transitions, behavior management talks, children waiting with nothing to do.
- *Mixed Content.* Not a code used by the observer. This code is used when more than one content area was chosen for one time-use episode, which indicates that the content being taught includes a 'mix' of subjects.

We combined time allocated to reading, language arts, and literacy as one content area – English Language Arts (ELA). Additionally, we combined time devoted to science and social studies. Finally, time observed on music/movement and art was combined into one content area.

Online Appendix D: Reported Time-Use (Pre-K Teacher Survey) Details

Part of our data collection included administering a survey to the pre-K teachers in our study. One of the items asked teachers to indicate how they allocated time during a typical school day. This time-use item, tailored to class-type, instructed teachers to report how much time they spent on each of ten listed subjects: mathematics, reading and language arts, social studies or science, eating, napping or quiet resting, art/music/drama/dance, unstructured play, structured play, and transitioning between activities. See Figures D1 and D2 for an example of the survey item by cohort. To reduce survey fatigue, we only asked about half-day teachers' AM classes in the surveys and assumed that they would describe time use for their PM classes the same. We have chosen to leave missing teacher surveys out of the analysis. Tables D1 and D2 present descriptive statistics by cohort, class-type, and subject for reported time-use collected from the teacher survey in mean minutes per day.

Table D1. Descriptive statistics for Cohort 2 by class-type and subject from the teacher survey in mean minutes per day

	Half Day (AM & PM) $n = 8$				Full Day $n = 8$			
Teacher Survey Subject	Mean	SD	p25	p75	Mean	SD	p25	p75
ELA	16.3	6.7	10.0	22.5	37.8	12.8	28.1	45.6
Math	13.8	4.4	10.0	17.5	25.6	12.0	17.5	35.0
Science & Soc. Stud.	12.2	2.4	10.0	14.4	20.2	16.5	12.5	18.8
Music and Visual Arts	12.2	5.5	7.5	16.9	15.8	3.9	13.1	20.0
Struct. Play	16.3	4.4	12.5	20.0	31.6	15.5	20.0	41.3
Unstruct. Play	67.5	17.1	57.5	77.5	97.2	29.5	72.5	120.0
Transitions	15.9	4.2	12.5	19.4	19.7	6.6	15.0	23.8
Eating	18.8	4.4	15.0	22.5	51.6	12.3	42.5	60.0
Napping	0.0	0.0	0.0	0.0	60.6	1.8	60.0	60.0

FN: Cell contents report the average number of minutes per day reported on each subject listed within the teacher survey. Results for half-day classes are shown on the left (N=8 total classes in Cohort 2), and full-day classes are shown on the right (N=8).

Table D2. Descriptive statistics for Cohort 3 by class-type and subject from the teacher survey in mean minutes per day

	Half Day (AM & PM) $n = 8$				Full Day $n = 9$			
Teacher Survey Subject	Mean	SD	p25	p75	Mean	SD	p25	p75
ELA	21.1	9.3	15.0	27.3	41.6	18.1	30.0	60.0
Math	22.7	12.6	14.4	31.0	25.9	11.0	20.0	30.0
Science & Soc. Stud.	11.3	2.3	10.0	12.5	23.6	10.0	15.0	30.0
Music and Visual Arts	11.9	5.8	7.5	16.3	26.1	6.0	20.0	30.0
Struct. Play	26.6	13.4	15.6	37.5	35.8	6.8	30.0	40.0
Unstruct. Play	49.0	30.3	30.5	67.5	104.3	22.8	90.0	120.0
Transitions	13.8	5.8	10.0	17.5	31.1	13.6	30.0	30.0
Eating	18.8	2.3	17.5	20.0	50.0	19.2	40.0	60.0
Napping	0.0	0.0	0.0	0.0	65.6	11.3	60.0	60.0

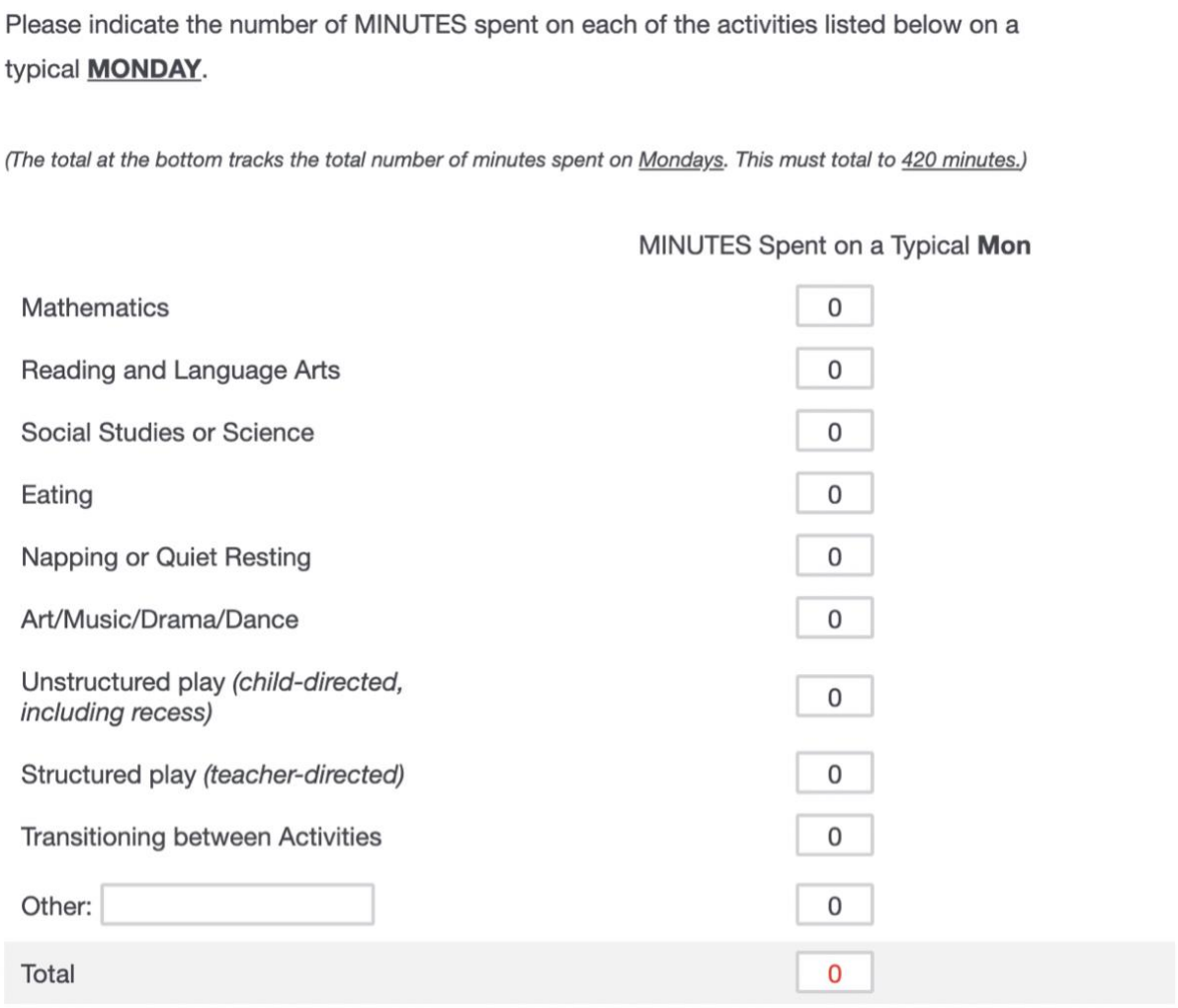
FN: Cell contents report the average number of minutes per day reported on each subject listed within the teacher survey. Results for half-day classes are shown on the left (N=8 total classes in Cohort 2), and full-day classes are shown on the right (N=9).

Please indicate the number of MINUTES spent on each of the activities listed below on a typical **MONDAY**.

(The total at the bottom tracks the total number of minutes spent on Mondays. This must total to 360 minutes.)

	MINUTES Spent on a Typical **Mon**
Mathematics	0
Reading and Language Arts	0
Social Studies or Science	0
Eating	0
Napping or Quiet Resting	0
Art/Music/Drama/Dance	0
Unstructured play *(child-directed, including recess)*	0
Structured play *(teacher-directed)*	0
Transitioning between Activities	0
Other:	0
Total	0

Figure D1. Cohort 2 Teacher Survey Time-Use Item for Monday



Please indicate the number of MINUTES spent on each of the activities listed below on a typical **MONDAY**.

(The total at the bottom tracks the total number of minutes spent on Mondays. This must total to 420 minutes.)

	MINUTES Spent on a Typical **Mon**
Mathematics	0
Reading and Language Arts	0
Social Studies or Science	0
Eating	0
Napping or Quiet Resting	0
Art/Music/Drama/Dance	0
Unstructured play *(child-directed, including recess)*	0
Structured play *(teacher-directed)*	0
Transitioning between Activities	0
Other:	0
Total	0

Figure D2. Cohort 3 Teacher Survey Time-Use Item for Monday

Online Appendix E: Cross-walking Procedure to Compare Content Areas across the Survey and Observation Protocol

To examine the consistency of time allocation across the teacher survey and the classroom observations, our team crosswalked the content areas and activities that appeared on each of these two instruments. First, we noticed that transitions and napping are directly comparable across the two approaches. We also code *eating* from the survey to be equivalent to the *meals and snacks* activity in the observation protocol as the language indicates the same activity is happening although the exact wording differs. Time-use episodes from classroom observation data labeled as either *reading*, *language arts*, or *literacy* have been combined and coded as *reading and language arts* to match the corresponding teacher survey item, although both of these have been relabeled as "ELA" for simplicity. Episodes that were coded either science or social studies as the primary content area were summed and coded as *social studies or science* since these activities were included as a single item on the teacher survey. Time-use episodes coded as *visual arts*, *specials*, or *music* were combined and included under the code *art/music/drama/dance* that was used on the teacher survey.

Overall, cross-walking academic subjects is not as straightforward as one may think if we believe that ECE teachers often include content instruction *within* time devoted to play. The survey includes two different categories for teachers to indicate how much time is spent on play in their classrooms – *structured play (teacher-directed)* and *unstructured play (child-directed, including recess)*. Thus, the common practice of incorporating academic content into playtime at this age level complicates cross-walking the separate academic content areas across the two data sources.

We also struggled directly comparing time devoted to play across the two approaches even without considering that academic content may be included in this time. We identified *gross motor*,

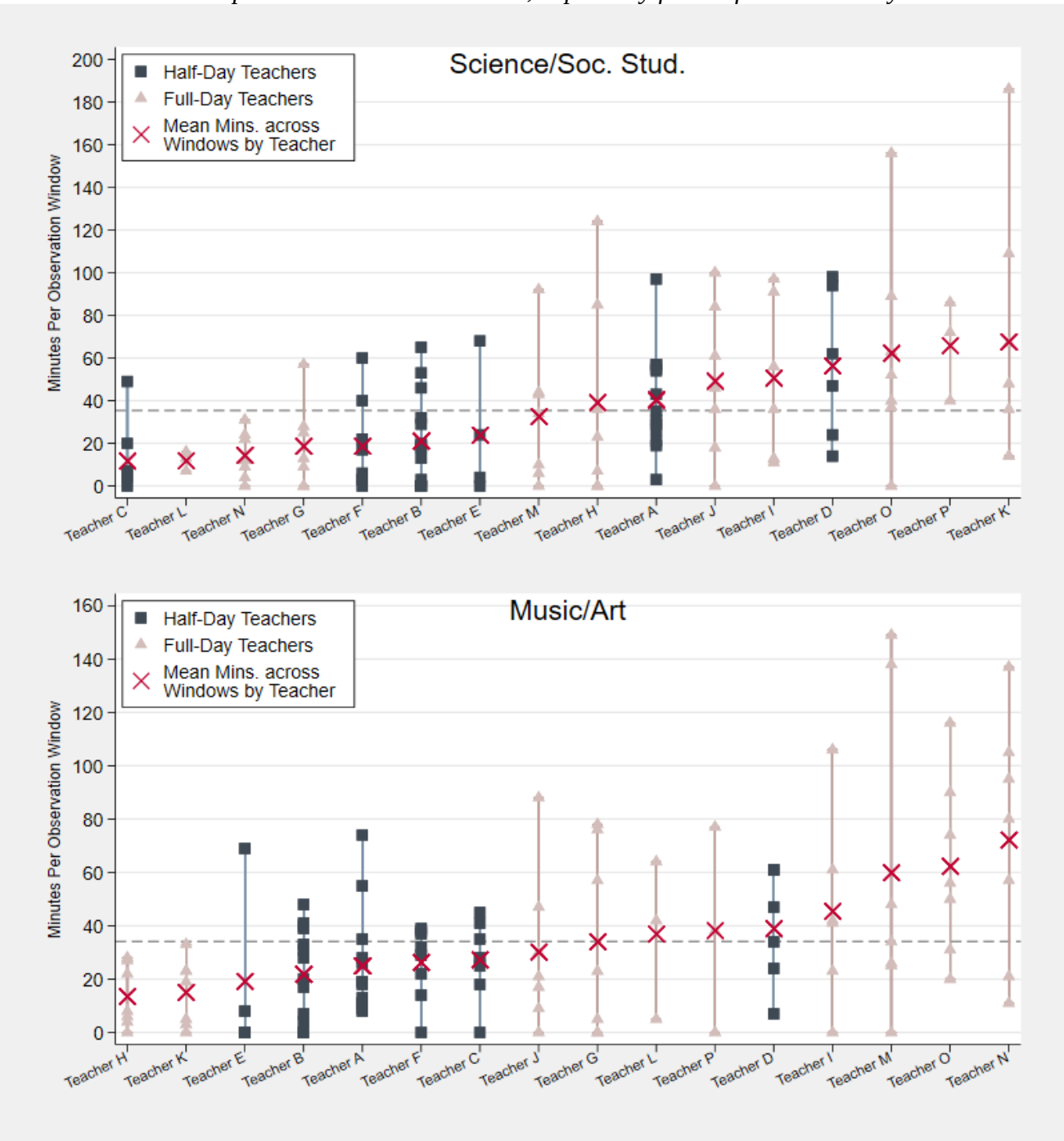
small group only, and *small group + centers* as codes that indicated the time-use episode in our observation protocol involved teacher-directed play and matched these episodes to the corresponding item on the teacher survey, *structured play*. It would be possible that combining time-use episodes across subject codes (*gross motor*) and activity codes (*small group only*) could result in 'double-counting' time; a single time-use episode could be coded for both the gross motor subject and a small group only activity type within the observation protocol and therefore the duration of the episode would be counted twice.

Additionally, we identified *playground* and *centers only* as codes from the observation protocol that indicated the time-use episode involved child-directed play or recess; therefore, time coded for either of these activity types within the observation protocol was combined and re-coded *unstructured play* to match the respective item on the teacher survey. Although the code *small group + centers* does suggest that at least some children are engaging in self-directed play, we chose to only include episodes with this code as *structured play* as the teacher assigned children to either different activities or to free-choice centers. Once again, we found that comparing play across the teacher survey and observational protocol to be challenging due to differences in these instruments' design.

Time-use episodes with the *whole group* activity code will already be counted in the previously mentioned content area codes, so we leave this activity code out of our present analysis since it does not have a potential matching category on the teacher survey. Episodes labeled either *none* or *content not applicable* will also be counted in the *downtime* activities (e.g. *transitions*).

Online Appendix F: Variability in Teachers' Time Allocation across Other Content Area and Activity Types

Figure F1. Science/Social Studies & Music/Art: Variation across Each Teacher's Observation Windows in Minutes Spent on Given Content Area, Separately for Half- and Full-Day Classes



Teachers (x-axis) were observed for an average of 6 complete class periods over the two-year time frame (min=3, max=14). Each dot represents an entire class-period observation. The red "×" on each teacher's vertical line represents the mean minutes that teacher spent across all her observations on the given content area. The dashed grey line indicates mean time spent on given content area overall (across all teachers and observation windows).

Figure F2. Whole Group, Small Groups Only, and Small Group/Centers: Variation across Each Teacher's Observation Windows in Minutes Spent on Given Activity

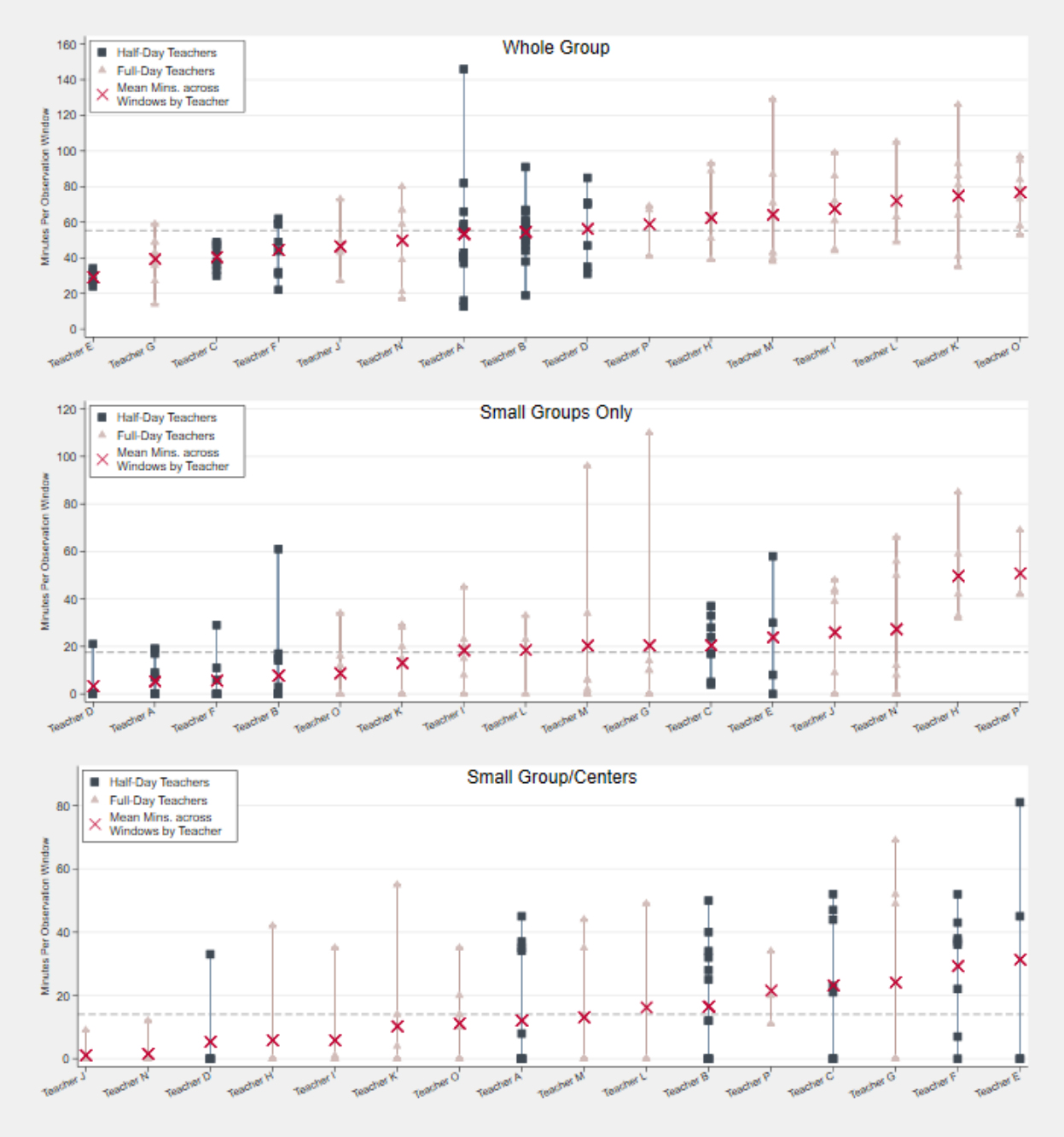


Figure F3. Centers Only, Transitions, and Eating: Variation across Each Teacher's Observation Windows in Minutes Spent on Given Activity

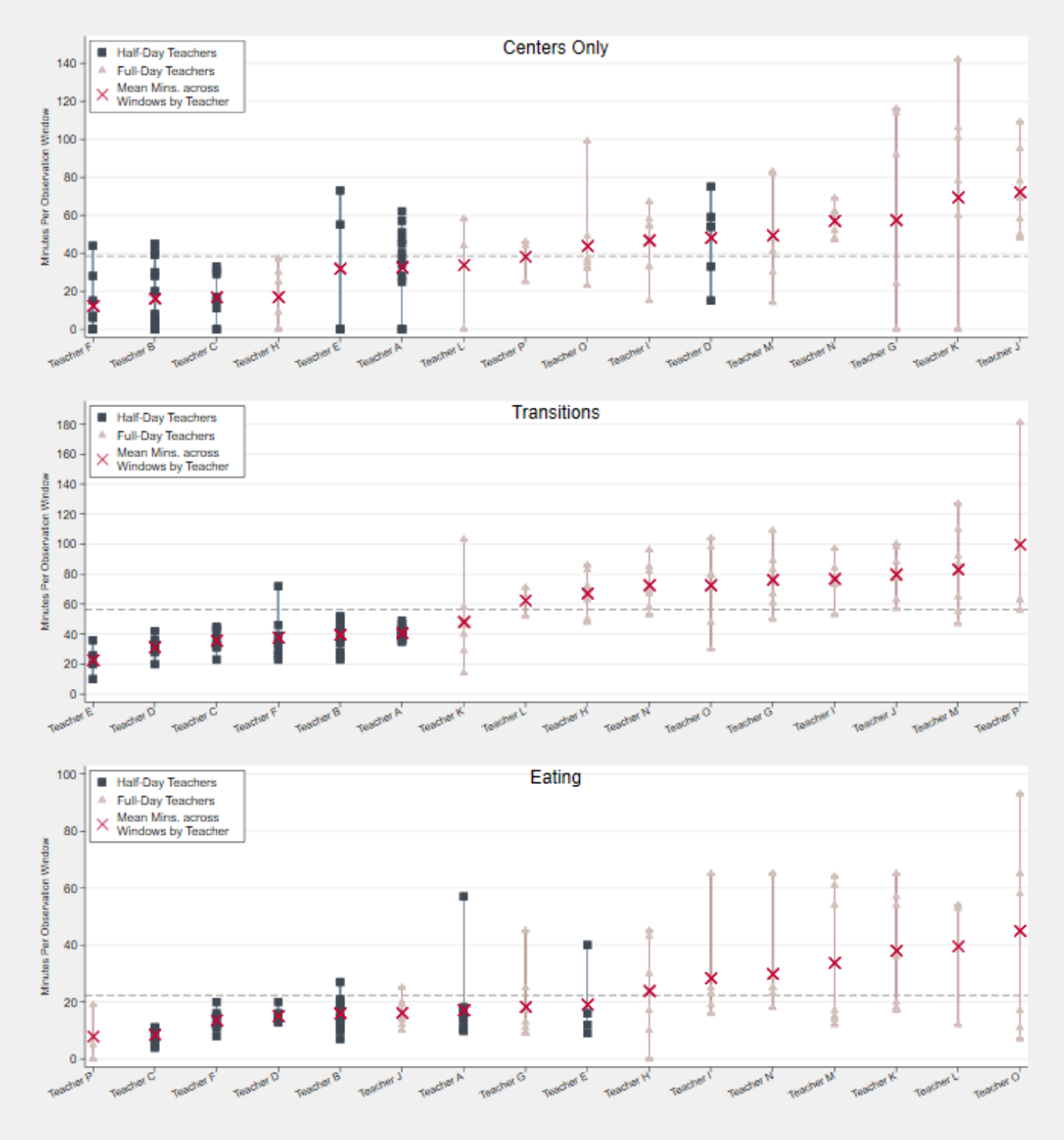


Figure F4. Napping, Playground, and Specials: Variation across Each Teacher's Observation Windows in Minutes Spent on Given Activity

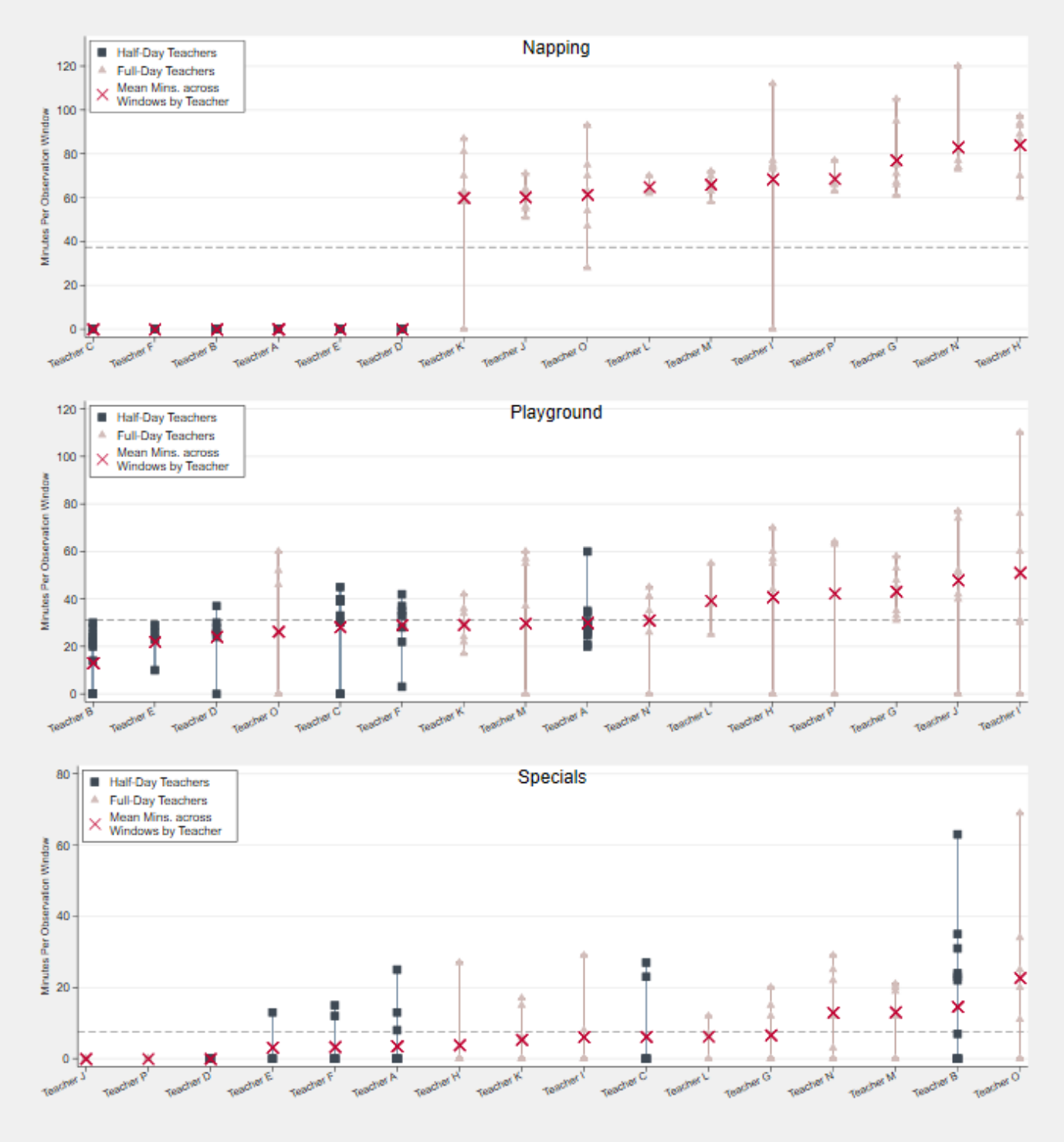


Figure F5. Variation across Teachers in Minutes Spent on Observed Activities for Half-Day Classes

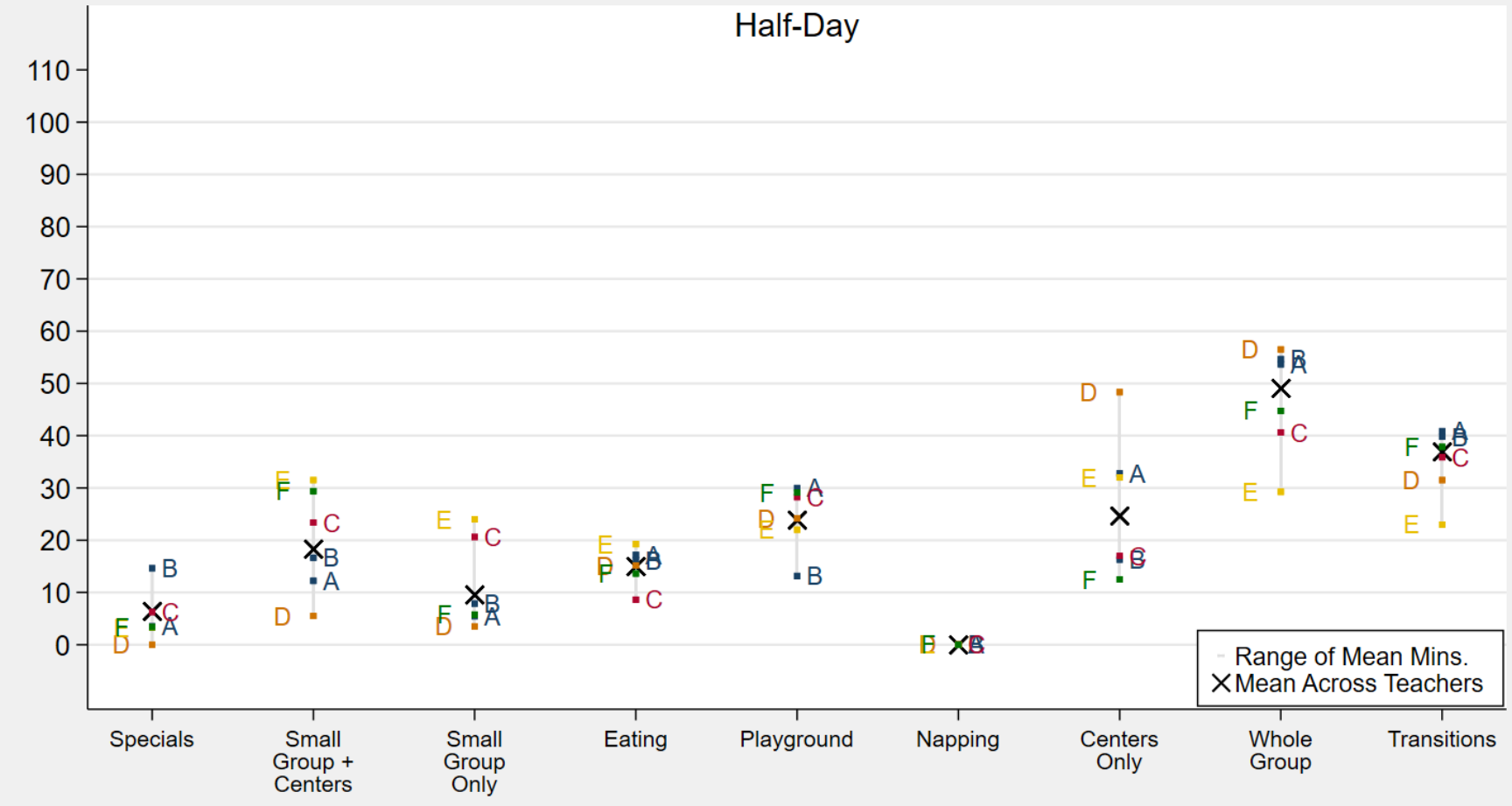
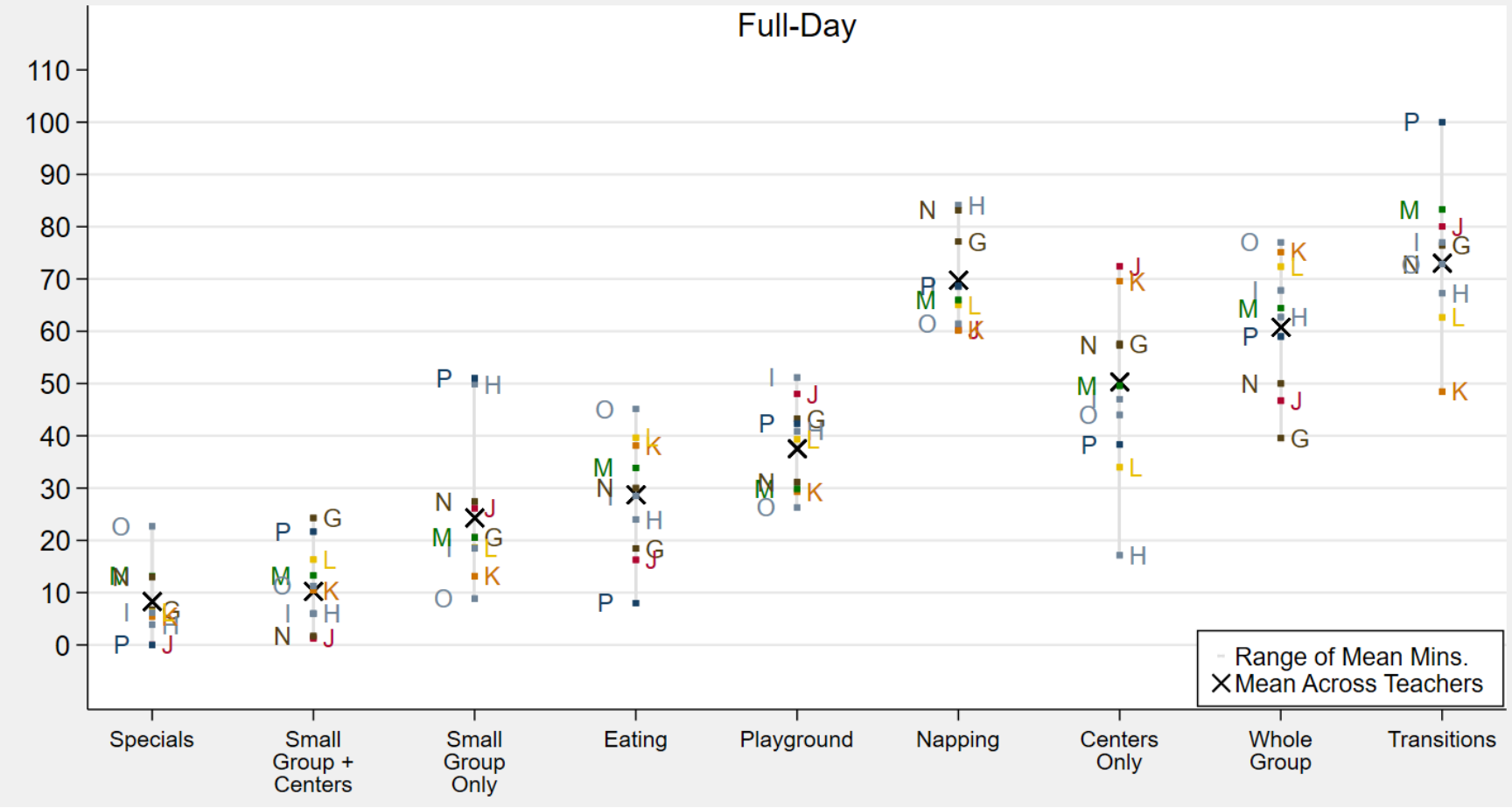


Figure F6. Variation across Teachers in Minutes Spent on Observed Activities for Full-Day Classes



Endnotes

[1] Outcome measures included the Peabody Picture Vocabulary Test (PPVT), the Teaching Strategies GOLD (TS GOLD), and the Early Screening Inventory-Revised (ESI-R). Author(s) found positive and statistically significant intent-to-treat (ITT) effects on the PPVT, as well as TS GOLD's cognition, literacy, math, physical, and socio-emotional development domains. While the analyses were underpowered due to the initial availability of only one cohort of data, the same study found positive, but as-yet, not significant effects on an end of pre-K ESI-R.

[2] For statistical power reasons, it was not feasible to design the FDPK Study's RCT to conduct a class-level mediation analysis of whether any full-day pre-K causal effects were mediated by class time-use (given that time-use data varies only at the classroom level for 18 treated classrooms in Cohorts 2 and 3).

[3] Large studies that collect classroom time-use data via teacher surveys data include the Early Childhood Longitudinal Study-Kindergarten 1998 and 2011 Cohorts (ECLS-K), the National Education Longitudinal Studies; the Schools and Staffing Survey; the Reform Up Close Study; and Study of Instructional Improvement.

[4] Data was collected for pre-K as a part of the NCEDL Multi-State Study of Pre-Kindergarten during the 2001-02 school year in six states: California, Georgia, Illinois, Kentucky, New York, and Ohio. Pre-K data collection for the SWEEP Study took place during the 2003-04 school year in five states: Massachusetts, New Jersey, Texas, Washington, and Wisconsin.

[5] Since the district rightly anticipated that more families would be interested in full-day pre-K than they would be able to accommodate, district leaders held lotteries to fairly allocate the limited full-day slots. This lottery process provided a unique opportunity to employ an RCT to estimate the causal effects of being offered/attending full-day pre-K on student outcomes, relative to being offered/attending the half-day program (Author(s)). The interested reader can find additional detail on the randomization in Author(s). Since interest remained high, WPS added two additional full-day classrooms for Cohorts 2 and 3.

[6] The current paper is analyzed at the class level; however, it may be useful to have a sense of the RCT child-level sample sizes: In Cohort 1, we randomized 226 applicants for 2016-17 pre-K to spots in either full- or half-day classes. Likewise, in Cohort 2, we randomized 264 applicants for 2017-18 pre-K. In Cohort 3, we randomized 310 applicants for 2018-19 pre-K.

[7] Curricula include: Little Treasures, a comprehensive curriculum available free online; Incredible Years, a social-emotional curriculum; and Hand Writing without Tears for fine and gross motor skills.

[8] Classroom observation data was only collected for pre-K classrooms in Cohorts 2 and 3. Half-day students did not have a lunch period (due to either leaving school before lunch or arriving at school after lunch). In full-day classrooms, we did not record time-use during the lunch/recess block.

[9]This analysis uses 34 unique pre-K class-years—17 classes in the first year (Cohort 2) and 17 classes in the second year (Cohort 3). There were 18 full-day class-years, 9 in each of the two cohorts.

[10] Classrooms in this district had both a lead teacher and assistants. We surveyed lead teachers only about how time was typically allocated.

[11] As a district, WPS requires all lead pre-K teachers to have earned at least an associate's degree and to have completed at least two courses in early childhood education as a part of their degree.

[12] We added a video component to the study in Cohort 3 to create a lasting data source for looking into the classrooms as future hypotheses emerged about classroom dynamics. Videos were made by a professional videographer using three simultaneous cameras and audio-recorders per class session.

[13] The possible activity codes included: (1) Meals & snacks, (2) Transitions, (3) Napping, (4) Playground, (5) Whole group, (6) Small group only (no centers), (7) Small group + centers, (8) Centers only (no small groups), and (9) Specials. For a more detailed description of these codes, see Online Appendix C.

[14] The possible content codes were: (1) Reading, language arts, & literacy (combined into one); (2) Math; (3) Science & social studies (combined into one); (4) Music and visual arts (combined into one); (5) Gross motor; and (6) Other. Observers could also select (7) "No content", or (8) "Not applicable". We added a code (9) for "missing content" for the rare cases in which the observer was not clear about the episode content. For a more detailed description of these content codes, see Online Appendix C.

[15] For example, if an observer coded a time episode that involved a shared book reading that involved conversations about being a good friend as whole group, they would also indicate that the activity included two content areas: Reading/Language Arts/Literature and Science/Social Studies. Thus, this example time-use episode would be coded as whole group for an activity and "mixed" for content. Mixed content is described in more detail below.

[16] For example, the observational protocol had separate categories for "reading", "language arts", and "literacy". These were combined into a single category in order to compare it to the "reading and language arts" category from the teacher survey.

[17] For instance, a teacher who teaches social studies only on Tuesdays would have the opportunity to indicate that their schedule on this day was different from other school days. After completing their Monday schedule, teachers were then allowed to indicate whether their schedules on Tuesday were the same as Monday or different. If they indicated that their schedules were different, they were given another item, identical to the Monday item, but asking about time-use on Tuesday. If they indicated that their Monday and Tuesday schedules were the same, their Monday response was saved for Tuesday, etc., for all days of the week, as applicable. See Online Appendix D for examples of the survey items used in this analysis.

[18] We included a data validation feature to the online survey, which required teachers to ensure the reported minutes spent in each category equaled the number of hours in that school day. However, Cohort 2 full-day teachers were mistakenly asked to total their survey time to equal a 5-hour school-day (instead of 6-hours). In Cohort 3, we corrected the full-day validation totals to reflect a 6-hour day, rather than a 5-hour day. There do not seem to be systematic responses from teachers in how they handled this problem (i.e., they did not systematically leave out an hour of eating or proportionally shrink reported time across all subjects).

[19] In the third cohort, the missing teacher had left the school district by the time the survey was administered. The response we received was instead completed by a separate person—the para-educator who had been in the classroom with those students the entire year and who had been responsible for teaching in small group activities. We therefore opted to not include this in a survey-observation comparative analysis since it was not time-use by the same person.

[20] While we use math as an illustrative example, in practice the analysis is conducted separately for each of the content/ activities that appear in the survey-observation crosswalk presented in Table 5.

[21] We had a total of 106 observations for our half-day analysis including 53 observations collected across 16 class-years matched with their respective teacher survey for that year (eight unique, half-day teacher surveys). For full-day, we used 108 observations in our analysis with 54 observations collected across 18 class-years matched with their respective teacher survey for that year (16 unique, full-day teacher surveys). We were missing two full-day teacher surveys and consequently excluded their classroom observations in the missing survey year from this analysis. Therefore, the Ns in this analysis reflect two sources of data multiplied by the 53 and 54 observations collected across class-years (half- and full-day, respectively) that could be matched to annual survey data collected at the classroom level.

[22] Whenever we present the percentage of the school day spent on a given activity or content area, we always do include time spent napping in the denominator of the full-day school day, unless otherwise noted. The regulation of children's sleep schedules may be supported by the afternoon nap could potentially play a role in the benefits of full-day. We therefore do not wish to exclude it.

[23] In the few instances where we refer to a percentage of the "awake school day", we have chosen to exclude the nap as part of that calculation. In all other cases, percentages of the school day count the nap as part of the school day.

[24] Of these, 42 minutes were "content not applicable" and 14 minutes were "no content". For a description of the difference between the "content not applicable" and the "no content" codes, see Online Appendix C.

[25] Some readers may be interested in additional detail about what activity types most often involved mixed content. Across all pre-K classrooms, we commonly observed mixed content in whole group activities, like morning meeting, in which teachers would switch topics quickly (e.g., counting days on the calendar to identifying the weather to singing a greeting song). Mixed content also frequently occurred during group-based activities (i.e., small groups, centers, and small groups + centers). We observed students either choosing from a number of different activities (centers) or rotating through a set of activities (small groups) that included a variety of content areas such as blocks, dramatic play, art, math games, sensory exploration (i.e., water tables), and literacy stations.

[26] In this analysis, we count all minutes spent on a given content area, even if it is being mixed in with another content area. In other words, the number of minutes in the day does not represent a zero-sum situation. Therefore, while there will always be trade-offs in time allocation, a teacher could combat this issue by finding ways to deliver multiple content areas simultaneously. Hypothetically, if a teacher could deliver all 5 content areas simultaneously in every minute of the day, she could spend the most time on *all* five content areas, relative to her colleagues. This is unlikely of course, and not consistent with patterns observed in Figure 3.