Supporting On-line Materials for 1139719

Study Overview

The current investigation uses observational data to report 1) descriptions of teacher and child behaviors and dimensions of quality of 5th grade classrooms; 2) teacher, school, family, and child correlates of these 5th grade classroom observations; 3) patterns and rates of exposure to stably high or low quality classroom settings across 1st, 3rd, and 5th grade; and 4) the extent to which poor children or those scoring in the lower third of the distribution on achievement at school entry are exposed to stably high or low quality elementary school experiences. This is the third in a series of reports on the nature of children's experiences in school and their influence on development for the approximately 1,000 children enrolled in the NICHD Study of Early Child Care and Youth Development (NICHD SECCYD) throughout the elementary school period. (NICHD: National Institute of Child Health and Human Development) This comprehensive, prospective study of a large sample of children in multiple sites across the United States affords a unique perspective on elementary education.

Method

Participants

The recruitment and selection of child participants in the NICHD SECCYD is described in several documents that are publicly available, to which interested readers are referred (http://secc.rti.org/). Families were recruited through hospital visits to mothers shortly after the birth of a child in 1991 at ten locations in the U.S. (Little Rock, AR; Irvine, CA; Lawrence, KS; Boston, MA; Philadelphia, PA; Pittsburgh, PA; Charlottesville, VA; Morganton, NC; Seattle, WA; Madison, WI). Of the initial pool of eligible mothers contacted for participation, 1,364 completed a home interview when the infant was 1 month old and became study participants.

The recruitment procedures were not designed to produce a probability, population sample. Below is described the manner in which we selected 5th-grade classrooms for observation, the procedures for which are identical to those followed in 1st and 3rd grades (*1, 2*).

In the children's sixth year of school (which for the majority was 5th grade), a total of 956 of the 1,364 original participants were observed in their classrooms; attrition reduced the participant pool to approximately 1,100 so that roughly 90% of retained participants were observed. From this pool, classrooms were then dropped if they were not a 5th-grade classroom, if there were fewer than 10 children present during the observation (less than 3% of the classrooms), or if the classrooms were otherwise atypical, resulting in a next-stage sample of 795 5th-grade classrooms. In 58 of these 795 classrooms, more than one study child was observed (always on different occasions by different observers). To provide a classroom-based focus for analysis, one child was chosen at random from those classrooms in which more than one study child was in attendance.

It is possible that the procedures for determining the observation period and circumstances imposed constraints that could affect the results. However, the procedures were designed to eliminate some outlier circumstances, such as a teacher with an atypically small number of children in her room or a self-contained special education classroom. The goal was to provide a picture of classrooms that would fit assumptions of what a typical U.S. classroom looks like. Furthermore, the procedures were designed to sample the classroom at its most instructionally-focused period, thus eliminating time devoted to activities such as assemblies, field trips, gym, recess or state-standards testing. The sampling constraints were intended to preempt a potential criticism that the observations were too inclusive, resulting in a picture of an under-stimulating classroom that was itself a function of the choice of when and what to observe.

The final sample consisted of 737 classrooms identified by the observer and teacher as a typical 5th -grade classroom in which 10 or more children were present at the time of the observation. It is important to emphasize that the 5th -grade classrooms observed in this study were selected only because a child enrolled in the ongoing prospective NICHD SECCYD was in attendance in those classrooms.   There were not enough schools in which multiple classrooms were located to support analyses of classrooms nested within schools; nor were there enough children per room to study nesting within classrooms. These 737 classrooms were distributed across more than 502 schools in more than 302 districts, and more than 33 states; these figures are low-end estimates because school and district information was obtained from the National Center for Education Statistics School and Staffing database and not all schools in the present sampled had information available. The majority of classrooms visited were located in and around the cities in which the 10 data-collection sites were also located. In the final sample of 737 classrooms attended by these children, 102 classrooms (13%) were located in private schools. An analysis that compared the private and public school classrooms on each of the teacher characteristics and observed classroom variables assessed in this study revealed only a few on which there were significant differences with no discernible pattern and fewer than would be expected by chance. Therefore the private and public samples were combined for analyses describing the classroom experiences to which this cohort of children was exposed as they attended 5th grade.

Teachers in these 737 third-grade classrooms were identified by themselves and the child's primary caregiver (usually the child's mother) as the primary teacher of that child. In team-taught classrooms a single, primary teacher was identified on the basis of having the most responsibility for the child's instructional program.   Table 1 presents information on

characteristics of the teachers and classroom settings observed during the school visits at 5th grade. Information for the same characteristics for the 1st and 3rd grade teachers included in a subset of analyses reported in this paper is available in two publications describing observations in 1st (*1*) and 3rd (*2*) grade classrooms and readers are referred to those sources for further details.

Of the entire sample of teachers, 99.5% percent had a bachelor's degree or higher level of education; 43.9% had either "some graduate courses" or a master's degree. These were experienced teachers; on average the teachers had over 12 years of teaching experience. For this sample of 5th grade teachers, the overwhelming majority (88.2%) met state standards for licensure and certification to teach in that state. The present sample included some private schools (see below) which account for some of the non-certified teachers. Teachers' ages and monthly salary ranges were quite variable. Teachers' self-reported ethnic status was as follows: Caucasian: 92.1%; African-American: 2.5%; Hispanic: 1.7%, other: 3.8%. In 5th grade, on average, child-teacher ratio was approximately 19:1 and there was considerable variability, with a range from roughly 5:1 to 36:1. Other adults, including paid and unpaid aides, parent volunteers, and student teachers, were typically not present during the observations.

The characteristics of the teachers and classrooms reported above for the NICHD SECCYD sample closely mirror national averages. For example, the mean age of teachers in the 1999-2000 Schools and Staffing Survey (*3*) was 42.3 and average years of teaching experience was 15.2. For the 1993-1994 year (*4*) full-time elementary school teachers, on average, had 15 years of teaching experience; 99% had at least a bachelor's degree, and 42% also held a master's degree; the average annual salary was $33,600. Thus, the present sample has somewhat less experience teaching, but educational levels and salaries closely approximate national averages.

Of the 737 study children observed in these visits, 143 (19.7%) were from single-parent homes. Ethnicity of the children was somewhat diverse, with 582 (79.0%) white, 75 (10.2%) black, 46 (6.2%) Hispanic, and 34 (4.6%) of other ethnic backgrounds. The average family income-to-needs ratio (based on US census definitions for poverty levels) for this sample was 4.68, with 19.9% of the sample having an income-to-needs ratio below 2.0. Mothers averaged 14.58 years of education, with 25.1% having a high school diploma or less.

To the extent that the NICHD SECCYD sample reflects a diverse range of family backgrounds as present in the United States, these observations reflect a typical day in 5th grade for a large number of children. However, because the NICHD SECCYD sample excluded children of adolescent mothers, mothers who did not speak English, and children who were hospitalized for more than a week at birth or who had a diagnosed disability, it is less inclusive of children experiencing a range of potential risk conditions. The sample included in the present investigation is somewhat more white and of higher income levels than nationally representative samples of families with children of similar ages (see *3*).

Measures

Observations took place during the spring of the child's 5th grade year between January and June. More than 75% of the observations were conducted during the 2001-2002 school year, and the remainder was conducted in the 2002-2003 school year.

Classroom Observations

Classroom observers used the Classroom Observation System for Fifth Grade (COS-5). The COS-5 is identical to the COS-3, used in the 3rd-grade classroom visits for this sample (*2*), which itself is an upward extension and extensive revision of the COS-1 (*1*). The focus of the observation is the classroom as well as a specific child and his or her experiences in the

classroom. It is important to note that the COS-5 observation and procedures were intended to describe selected experiences in 5th -grade classrooms for the specific children enrolled in the NICHD SECCYD; they were not designed or implemented because of any characteristics of the child, the child's family, or the school. Observations were recorded at two-levels: a) teacher interactions with the target child and the activities in which this target child was engaged and b) global dimensions of classroom interactions and processes to which the target child and all other children in the class were exposed.

All observations occurred during the morning and began with the official start of the school day. The entire observation took approximately 6 hours. A number of codes were used to index each of several broad categories of behavior: the target child's behavior, the teacher's behavior toward the child enrolled in the study, the setting in which the target child was working, and the overall classroom environment. Discrete child and teacher behaviors, specific activities, and setting conditions were coded on a time-sampled basis while global ratings of the classroom and teacher were made using a set of seven-point rating scales. A detailed manual describes these codes and the coding procedures (see http://secc.rti.org).

Time sampling. For the formal COS-5 observation, eight 25-minute observational cycles took place. In each cycle observers made time-sampled recordings of discrete codes for one 10-minute period comprised of paired 30-second "observe" and 30-second "record" intervals. Thus there were 80 different minutes in which discrete behaviors were sampled across the observations throughout the school day.

Time-sampled codes included measures of the <u>setting</u> (e.g. whole class, small group, individual), <u>activities</u> (e.g., literacy, math, science, social studies, transition/management), <u>teacher behavior</u> (e.g., attends to child, teaches basic skills, teacher analysis/inference,

managerial instruction, affect positive, affect negative, disciplines), and <u>child engagement</u> (e.g., whether the child was engaged in the assigned activity). Importantly, these codes reflect the settings, activities, and teacher behaviors to which all children were exposed (individually or in group settings). A more specific level of coding was also performed for any activity that involved a literacy or math component. When an activity involved a literacy component (e.g., a literacy lesson or a social studies lesson in which a book was being read), coders also recorded whether the activity was "word level" or "comprehension." Word-level activities in literacy referred to those in which the skill taught or performed involved within-word information such as decoding, learning letters, etc. Comprehension activities had a focus on deriving meaning from print, of either individual words (vocabulary lessons) or passages. Similarly, if an activity had a mathematics component (e.g., a math lesson or a science lesson involving calculation), coders noted whether the activity involved "computation" or "concept development/problem solving."

Within any category of codes (e.g., teacher behavior), the codes are not exhaustive. The codes within categories are mutually exclusive, although double-coding within an interval is allowed if two different codes are evident in that interval. In addition, as was noted earlier, the observation procedures limited the period of observation to times in which instruction was most likely to occur, eliminating recess, gym, assemblies and other explicitly non-academic time. Coders were also asked to note whether or not the teacher administered a test or quiz during any portion of the observation cycle; only 8% of cycles included administration of a test or quiz.

Global ratings. In addition to the time-sampled coding, observers were assigned five minutes before each time-sampling period, as well as 10 minutes at the conclusion of the time-sampled period, to conduct observations and take notes about the classroom environment.

Coders relied on these dedicated periods of observation, as well as what they observed during the "observe" intervals of time-sampling, to make global ratings of classroom quality and teacher and child behavior using a set of seven-point rating scales.

Classrooms were rated on nine different dimensions of emotional and instructional climate using 7-point scales (*1, 2*) validated in prior studies (*5, 6*). Higher scores on these ratings were predictive of gains on standardized tests of reading and math in pre-kindergarten (*6*); reading in 1st grade (*5*) and math in 5th grade (*7*). Effect sizes for the validity of these scales predicting to achievement gains, adjusted for family process and demographic factors, are modest, in the .10-20 range across these various studies.

Global ratings of *classroom-level* dimensions included over-control, chaos, positive emotional climate, negative emotional climate, detachment of the teacher, teacher sensitivity, productive use of instructional time, and richness of instructional methods. These classroom-level ratings ranged from 1-7. A rating of "1" was assigned when that code was "uncharacteristic;" a "3" was assigned when the description was "minimally characteristic;" a "5" was assigned when the description of the code was "very characteristic" of the classroom; and a "7" was assigned under circumstances in which the code was "extremely characteristic" of the observed classroom. The global codes at the "high" end of the rating scale(s) reflect the extent to which the classroom or teacher demonstrated that particular dimension to an exceptional degree.

Teacher and Classroom Characteristics

Teachers completed a questionnaire providing information on their years of teaching experience, monthly salary, extent of post-high school education, involvement in in-service training activities during the last year, and their perception of principal support. Items for this

questionnaire were selected from the School and Staffing Survey Teacher Questionnaire, used as a standard part of the Common Core of Data collected on schools through the National Center on Education Statistics (*3*). Teachers also completed the Teaching Self-Efficacy scale (*8*), which provides a reliable and valid indicator of a teacher's beliefs and feelings of efficacy related to aspects of her role as a teacher. The Total Efficacy score was used in this study; Cronbach's alpha for this scale was .91. These teacher and classroom characteristics provided information related to the correlates of observed experiences in 5th grade. Children reported on the quality of the classroom environment in 5th grade using a self-report measure of school bonding and relationships (*9*)

Child and Family Characteristics

Mothers reported on levels of family income both at 4th grade and prior to the child's entering kindergarten. Using federal guidelines for poverty levels, we computed an *income-to-needs ratio* at the time the child entered school and at 4th grade based on these maternal reports and information on household size. Information was also available on the child's level of cognitive ability at 4th grade, using the total score from the Wechsler Abbreviated Scale of Intelligence (*10*) and the child's 4th-grade teacher completed the Social Skills Rating Scale (SSRS: *11*) and the Teacher Report Form (*12*) at the end of the 4th grade year. Both of these instruments provided estimates of the child's functioning in the year prior to 5th grade that were used to examine sorting of children into 5th grade on the basis of prior-year functioning. Children also were assessed prior to the start of kindergarten, at 54 months, using the Woodcock Johnson Psychoeducational Battery-Revised (*13*), with the Word Attack and Applied Problems subtests used as a composite to reflect academic achievement skills prior to school entry.

Training and Reliability

Observers from all 10 sites trained on practice videotapes using a standardized manual that provided extensive descriptions of codes and anchor points, prior to attending a centralized training workshop. After the central training workshop coders returned to their sites, conducted pilot observations, and trained on two more videotaped cases. Following this training regimen, all observers had to pass a videotaped reliability test involving six cycles of time-sampled coding and qualitative ratings. Criteria for passing were at least a 60% match with a master coder on time-sampled codes and an 80% match (within one scale point) on the global rating scales. All coders passed at these levels on a reliability test before being certified to conduct observations in the field.

Average exact agreement with the master-coded videotape test for the time-sampled codes, estimated by correlation with master-coders' scores, was .85. For the global ratings, reliability was estimated as within one scale point on the seven-point rating scales. Average reliability for the child and classroom global ratings on the videotaped test was .79, again estimated by correlation with master coders' ratings. Observers each also conducted a minimum of two paired visits scheduled randomly during the data collection window for the purposes of estimating live reliability. Correlations between observers exceeded .84 for all but one of the 38 time-sampled codes, with the lower estimate (.62) due to the infrequency of the observed behavior. Average live reliability across all global ratings, estimated using correlation, was .71

As a final check on the reliability of the observational data, we examined the associations between codes for classrooms observed more than once. That is, because some classrooms were attended by more than one study child, we had a sample of classrooms (N = 54) in which we could examine stability of the observations across days/different children, a form of test-retest

reliability. For the time-sampled codes, the average cross-day correlation was 0.83 (SD = 0.09). For the global qualitative ratings of the classroom environment, the average cross-day correlation was 0.88 (SD = 0.10). Thus it appears that the one-day observations reported for these 5th grade classrooms were reflective of aspects of the classroom setting that remained stable across days (and different children). We found similar results for the 1st grade observations when the average cross-day correlation (N = 63 classrooms) for the time-sampled codes was .79 (SD = .15) and .71 (SD = .30) for the qualitative ratings; as well for the 3rd-grade observations when the average cross-day correlation (N = 52) was .87 (SD = .08) for the time-sampled codes and .91 (SD = .08) for the qualitative ratings. In short, it appeared these observations reflected stable aspects of the classroom setting.

Scheduling

After all permissions had been secured, observers contacted teachers to schedule the observation. During this contact, observers queried the teachers about the daily routines and schedule of the classroom. Using this information, the protocol required observers to schedule observation periods that would maximize the likelihood of observing instruction and minimize the number of nonacademic activities that would be observed.

Results

Results are presented in several sections pertaining to the research topics outlined earlier: descriptions of experiences and quality in 5th grade classrooms; associations with teacher, school, classroom, and child characteristics; and patterns of stability in classroom experiences across the elementary school period.

Classroom Settings, Activities, and Teacher Behaviors in 5th Grade Classrooms

The first set of results describes the setting in which the child was working, the activities to which the child was exposed, and the behaviors of the teachers in these classrooms. These results are based on time-sampled discrete behavioral codes recorded for the 737 5th grade observations that included a minimum of 80 intervals of recording distributed across approximately 6 hours. Note again that these observations were designed to capture the target child's experience in classrooms. These data were re-scaled to a 60-interval base. Table 2 displays descriptive information in two forms: the left-hand set of columns displays the number of intervals in which a code was observed (with information on the range), and the right-hand column presents the average percent of intervals in which that code was recorded. It is important to emphasize that these codes pertain to the activity and subject area to which time was devoted and the child was exposed during that specific interval. They are not exhaustive of the possible activities or subject areas that could, or might have been, offered to children. Also, these codes do not reflect whether or not the child was actually engaged in that activity. With regard to the possibility that children were taught by different teachers, these codes reflect experience across all teachers observed for the minimum of 8 observation cycles. In short, these findings should be viewed from the perspective of exposures for a typical child across the course of a day in 5th grade.

In terms of the setting in which the child was working, Table 2 shows that in the average classroom, a child was most often working in a whole group (52.8%) or individual seatwork (38.4%) settings. Together, whole group or individual settings accounted for 91.2% of intervals observed over the day. With regard to the specific subject area of an activity, in more than one-third of the intervals (36.7%) children were exposed to a literacy activity. Also occurring fairly

often were math activities (24.4%) and transition/management (17.1%). Science and social studies activities occurred in roughly 10% of the intervals, even though the coding system allowed that if an activity had literacy and social studies components, for example, then both literacy and social studies could be coded. Within the categories of literacy and math activities, comprehension appeared to occur somewhat more often (77.1%) than word-level work in literacy while computation dominated mathematics activities (71.2%).

It is important to note that for each type of activity the percentage of intervals in which children were exposed to that activity varied widely across classrooms. During these day-long observations, some students received very little exposure to a particular type of activity, whereas for others nearly all the time was devoted to that same activity.

Table 2 also describes the behaviors of the classroom teacher, aggregating across different teachers when necessary, again to provide a picture of a typical child's experience with a teacher. Overall, an individual target child was attended to by the teacher in 8.4 % of intervals, regardless of the group setting in which it occurred. The code of "contact with a study child" refers to any interval in which the teacher interacts verbally or non-verbally with the child in any type of setting. The results imply that children must manage their own learning and activity without contact with the teacher, for the vast majority of time. With regard to actual teaching behaviors, children were exposed to teachers' instruction in basic skills or analysis/inference skills, on average, in 38.2% of the intervals, with a ratio of teaching basic skills to analysis/inference of approximately 5:1. Teachers were also fairly often engaged in offering instructions on how to manage materials or time (18%).

Global Ratings of the Classroom Environment

Table 3 presents descriptive results for the global ratings (using 7-point scales) of the 5th grade classroom environment. For these results, the appropriate frame of reference is the classroom—these ratings were recorded based on teacher behavior and classroom-level interactions and qualities regardless of what the target child was doing. By and large these results indicate that observers had moderately positive impressions of the social aspects of the classroom setting, with average teacher sensitivity and positive classroom climate ratings of 4.85 and 5.14 (respectively) on these seven-point scales. These classrooms also appeared to be fairly busy settings in which children and teachers did have assignments and tasks they were performing, with a rating on productive use of time of 4.9 on a seven-point scale. However, in terms of features of the instructional environment, ratings were much lower. "Richness of instructional methods", which reflects the variety of formats used for instruction and the extent of conceptual focus, received a rating of 3.6 while "evaluative feedback," which reflects teachers' provision of frequent and direct feedback on student performance, averaged a rating of 3.44. These scales were also observed to show somewhat more variability across classrooms than the other dimensions. This pattern of ratings for instructional features reflects that although children were engaged in instructional activities, for the most part, they were exposed to only one method or mode of instruction (such as a vocabulary worksheet or watching the teacher do math problems on the board) and received fairly generic feedback regarding their performance that focused primarily on correctness. Using the scale-point descriptions for these dimensions, ratings of "3" indicate that aspects of the classroom environment to be "minimally characteristic" of the setting.

What Are Children Doing in 5th Grade?

A set of time-sampled codes was gathered on the child's behaviors during the 8 observational periods throughout the day. The descriptive findings for these codes are reported in Table 4. These results indicate that, of the opportunities for children to be engaged in a learning activity during the sampled intervals, children were at least minimally engaged (68.6% overall) more than twice as often as they were unproductive or not engaged. We emphasize that the threshold for this code is quite low (e.g. the child may be looking at the teacher during instruction) and does not measure active learning. Reflecting the nature of the teachers' instructional behaviors, children were mostly offered instruction that focused on basic skills, in contrast to analysis/inference or problem-solving, at a ratio of about 4.5:1. Collaborative work with peers such as peer-mediated cooperative learning, occurred quite rarely (4.8% of intervals) although children did have a fair number of positive/neutral interactions with peers; even more rare were negative behaviors with peers or affective displays toward the teacher (positive or negative). As was the case for classroom and teacher codes, these child behavior codes varied considerably across children.

Associations of Classroom Features with Teacher, Family, and Child Characteristics

We then examined relations between the two primary factor-based composites derived from the global rating scales (i.e., positive social climate and classroom instructional quality) and an assortment of teacher, student, and family factors, such as teachers' total years of public school teaching experience and their salary, family income, and child ability and skills assessed in prior years. Because nearly all the teachers in this sample were fully credentialed in their state, it was not possible to examine teacher credentialing. These correlations are reported in Table 5 and pertain to three questions: Do features of teachers and schools co-vary with

observations of the classroom? Are children with certain demographic attributes or skills as they complete 4th grade more likely to be exposed to higher or lower quality classroom experiences in 5th grade? Do these students' perceptions of their 5th grade classrooms relate to what was observed?

With regard to teacher characteristics, the classroom social climate was observed to be more positive when teachers reported having a greater sense of influence on school policy and reported more feelings of efficacy in the classroom. Higher class size was associated with somewhat lower observed levels of positive social climate. Classroom instructional quality was observed to be higher when teachers had somewhat fewer years of teaching experience, were paid higher salaries and reported more influence on school policy and more efficacy. It is important to note the magnitude of all of these associations is small and the causal direction is unknown. When we examined the collective predictive association of observed classroom quality by these teacher, classroom, and school characteristics (using regression) the variance accounted for (R^2) was 0.04 for positive social climate and 0.06 for classroom instructional quality.

Family characteristics appeared to have a somewhat stronger association with observed classroom quality; both social and instructional quality of the classroom were observed to be higher for children whose mothers had higher levels of education and whose families had higher income levels relative to their needs. As was the case for the school, teacher, and classroom correlates just described, the associations between family education and income levels and observed classroom quality, albeit significant, were small in magnitude.

We then examined the extent to which observed experiences in 5th grade classrooms covaried with attributes and perceptions of the child. Child gender was unrelated to observed classroom quality. Children rated by their 4th grade teachers as higher on externalizing problems

tended to attend 5th grade classroom rated as somewhat lower on social climate, relative to children without prior externalizing problems. Children rated as higher on social skills in 4th grade and those whose assessed cognitive ability in 4th grade was higher were exposed in 5th grade to classrooms rated higher on both social climate and instructional quality. These correlations are again fairly small in magnitude, and may suggest either selection/sorting effects (more capable children are placed in better quality classrooms the next year) and/or a child effect on the 5th grade ratings such that more capable children play a role in eliciting or creating higher quality classroom environments. However, readers are reminded that analyses reported earlier showed high levels of stability in the classroom ratings even across different days or when different children were the target. Finally, small, significant associations indicated that students had more positive perceptions of classrooms that were rated as higher on emotional and instructional quality.

Stability of Observed Experiences across Grades 1, 3, and 5

As noted earlier, all children enrolled in the sample were eligible to be observed in their 1st, 3rd, and 5th grade classroom placements, with results describing those experiences reported in other publications for 1st and 3rd grade (*1, 2*) and above for 5th grade. One of the goals of the present study was also to examine consistency in, or stability of, children's classroom experiences in these grades in terms of the global ratings as well as the time-sampled codes common to the observational systems used at each time. Note the focus of these descriptive analyses of stability is the *child*; the goal is to characterize consistency (or lack thereof) in children's classroom experiences across these 3 grades.

In one set of analyses we examined consistency at the most global level, in terms of the emotional and instructional dimensions of the classroom environment using the composites

created from the 7-point qualitative ratings conducted at each grade level—Emotional Quality and Instructional Quality. Each of these composites was created on the basis of factor analyses of a set of global ratings, and although the nature of these ratings shifted slightly over time to accommodate developmental changes as well as additional focus on instruction in later grades, factors reflecting these two dimensions were found consistently at each grade and several common scales were among the highest loading at each grade (see *14, 2*). Thus these composites reflect global aspects of the classroom that could be justified as similar across grades.

Table 6 presents the set of correlations among the emotional and instructional composites across- and within-grades. Importantly, this table presents this information for observations for the whole sample of children below the diagonal, which includes children who may have changed schools, as well as for the children who remained in the same school throughout the grade 1-5 period (above the diagonal). For the whole sample, the cross-grade correlations indicate there was little consistency in the quality of the classrooms in which children were enrolled across these years. Although the quality of the emotional environment was modestly stable from one observation to the next (correlations of r =.17 and r = .25), the quality of the instructional environment appeared more variable across time, with correlations ranging from r =.05 to r = .12. This pattern, and the magnitude of stability, was quite similar for children enrolled in the same school, with little evidence of substantially greater stability if children were enrolled in the same school. In general, these correlational analyses indicate that across the elementary school years, individual differences in children's experiences in classrooms vary considerably, even at the quite global level of analysis reflected in these composites of instructional and emotional quality.

Patterns of Stability in Classroom Experience

As a further analysis of stability, we examined the extent to which children were enrolled consistently in high or low quality classrooms using percentile-based cutoffs for determination of "high" or "low," based on the Emotional and Instructional quality composites at each grade. In these analyses, the distributions of the Emotional and Instructional quality composites were divided into terciles at each grade to characterize Emotional or Instructional quality as "high" (> 66th percentile) or low ($\leq$ 33rd percentile). Children were in "consistently" high or low quality classrooms if ratings on either of these composites were in the high or low tercile on 2 or more occasions across grades 1, 3 and 5 and if not in a high or low group on all 3 occasions, the rating on the third occasion was in the middle tercile. These results complement the correlational findings presented above and involve 994 children because we included all available children for whom observations were made on each occasion, rather than a one-child-per classroom-sample.

For exposure to emotional quality, 17% of children were enrolled consistently in classrooms rated in the top tercile (N= 169) while 19% were exposed regularly to classrooms rated in the bottom tercile (N = 184). With regard to instructional quality, only 14% (N = 141) of children were exposed to classrooms rated in the top tercile (note these are *high quality only as relative to the sample distribution)* while 20% were enrolled consistently in classrooms at the bottom tercile of instructional quality (N = 199). Sixty-six (7%) of children were enrolled consistently in classrooms rated as high on both instructional and emotional quality while 85 (9%) were placed at all 3 grades in classrooms in the lowest tercile on both composites.

In short, based on stability coefficients as well as patterns of relatively high and low quality, both inconsistency as well as mediocrity of experiences of classroom quality appears to be the norm; it should be emphasized that the quality cutoffs were determined relative to the

sample distribution of quality, not in terms of absolute ratings on these seven point scales. Were we to have divided the distributions on quality according to the markers for "characteristically high" quality defined on the scales, the rates of stable high quality would have been much lower (particularly for instructional quality).

Do Low-Income or Low-Performing Children Receive Exposure to Higher Quality Experiences?

As a final descriptive analysis related to stability in classroom quality across the 1st, 3rd, and 5th grades, we examined the extent to which poor children or low-achieving children (assessed at the start of school) were exposed to consistently high or low quality experiences in elementary school using the patterns and groupings of high and low quality described above. This analysis examines factors related to selection into patterns of stably high or low classroom quality and are qualified by the fairly small percentage of the sample that was enrolled in one of these patterns (see above). For achievement groupings we used scores from the reading and math scales of the Woodcock-Johnson collected at 54 months, composited these into a broad achievement score, and again divided the distribution into achievement terciles. Children were also grouped into 3 levels of income-to needs: poor, near-poor, and not-poor. Table 7 presents results for the overlap in classifications of high and low achievement with stable high and low classroom emotional quality and instructional quality. Note that these analyses involve a subset of the 994 children involved in the descriptions of stability just reported; in these analyses we used only those children who were exposed to consistent high or low quality classrooms. Results indicate that by and large the likelihood of exposure to stably high classroom emotional quality or classroom instructional quality throughout elementary school is higher for children entering school already high on achievement or not poor and lower for those achieving at the lowest levels or coming from poor or near-poor families.

Discussion

Detailed observations of a typical day for a typical child in their 5th grade classroom reveal a pattern of experiences very similar to those reported in observations at 1st and 3rd grade for this sample of children (*1, 2*) and for 5th graders in schools in Arizona (*15*)-- high levels of variability across classrooms in activities and experiences to which children are exposed, with the average day dominated by basic skills activities in whole group or individualized seatwork settings. Similar to the earlier grades, the social climates of 5th-grade classrooms are, on average, mildly positive, although instruction appears lower in quality. Cooperative learning activities, or activities in technology, social studies or science, are relatively rare compared to reading, math, or transitions. On average, the nature and quality of children's experiences in fifth grade appear nearly unrelated to the nature and quality of experiences to which they were exposed in first and third grades, and fifth grade experiences show only very modest associations with teacher attributes such as experience or classroom attributes such as child-teacher ratio. Family education and income levels and the child's competence in 4th grade show small positive associations with higher observed quality in 5th grade classrooms, although these associations account for very small proportions of variability in 5th grade experiences. In general children are not exposed to either consistently high or low quality classrooms across the 1st 3rd and 5th grades except when they enter school already scoring low (or high) on achievement, or if they are poor; thus there appears to be a notable selection effect for exposure to stably high or low quality. These results confirm prior results from this study (*1, 2*) and other work (*15, 16*) and have implications for national education legislation mandating high quality programming and

instruction, particularly definitions and assessments of high quality teaching and mechanisms for ensuring equal access to high quality educational experiences (*17-21*).

Once again our observations of a typical day in an elementary classroom setting in the United States, drawing from a sample of 737 classrooms, in more than 302 school districts and 33 states, confirm earlier reports of studies focusing on fewer classrooms, on different grade-levels in elementary school, or using different observational systems (e.g., *22-24, 1, 2, 25, 26*), in that early elementary classrooms vary widely in the nature and form of experiences offered to children. In the present study, we expand the heretofore cross-sectional perspective on variability and examined variation in children's experience across 3 grades in elementary school and found the same overall pattern with respect to the nature and quality of classroom experience: it is exceptionally variable. The bases for this variability were not obvious, in that there were not even moderate associations between these observations and structural features of the class (e.g., size or ratio) or teacher (e.g., experience or in-service training), with the exception being that children from households lower in income or educational levels, or children who were reported to be more poorly adjusted in the prior grade tended to be enrolled in 5th grade classrooms of somewhat lower quality.

Although few children experienced stably high or low quality across the 3 elementary grades studied, those that experienced stably poor quality classrooms (a higher percentage than for stably high quality) were precisely those children who needed high quality—they were disproportionately drawn from the bottom third of the distribution in achievement or social skills assessed prior to starting school. The "inequality at the starting gate" (*20*) appears to apply not only to differences in children's skills but in the opportunities to which they are exposed.

Although again the tendency toward "tracks" of high and low quality as a function of family or child characteristics is only slight and tempered by the exceptionally high levels of instability and inconsistency. In fact, it is of interest that in other investigations that appear to have operationalized "effective teachers" in terms of links to improved child outcomes (*27*) having such teachers for several years in a row is of considerable importance in maintaining gains attributable to a year in a good classroom—the findings in this study suggest consecutive years of high quality teaching is neither likely, nor easily predicted.

We emphasize that there is evidence that the aspects of classroom quality observed in this study—emotional support and instructional support—have been shown in first grade classrooms to contribute to the elimination of the achievement gap between risk and non-risk groups, when risk has been defined in terms of having a mother with less than a college degree or being identified as functioning poorly in the prior grade (*25*). Moreover, these dimensions have also been shown to predict growth in children's functioning in large-scale studies of pre-kindergarten children and classrooms (*6*), are related to variation in achievement trajectories (*28*) and children's social functioning (*7*) in the present sample, and are closely related to observed aspects of classrooms that other investigations have shown to be related to positive child outcomes (e.g., *22, 23, 29, 30*). In short, the observations described herein, for which variation appears to be so pronounced, also appear to be aspects of classroom resources that predict children's achievement and social functioning, net of family influences and prior performance. To the extent that there have been observational studies in large samples of the "value-added" dimensions of what actually takes place in classrooms, these observations appear to capture at least a portion of that effect. Thus, when children's exposure to such resources, within or across

grades in elementary school appears to be virtually unrelated to structural features of schooling used to regulate quality (*18, 21*) and when such resources are deployed such that the most needy children are less likely to be recipients, it should be of considerable concern to educators and policymakers interested in ensuring equal access to high quality teaching (*17*). In fact, these data suggest that the definitions of a "high quality teacher" that currently dominate teacher licensure, training, and certification in the states, could be somewhat suspect—these data suggest that a child could be exposed to a teacher who meets NCLB requirements for high quality and yet observations of that child's experience in that teacher's class might lead to other conclusions.

That variation in classroom quality is not a function of class or teacher attributes typically used to regulate teacher quality or ensure some degree of uniformity in children's experience or instruction (*31, 32*) raises questions about what factors should be the appropriate focus of regulation, and whether attention to these parameters as the keys to raising child outcomes and as the indicators of quality (*33*) is somewhat misplaced. In this respect these results are not inconsistent with analysis of Texas data showing little to no effect of teacher education or experience on student achievement (*34*), analysis of elementary schools in North Carolina in which teacher experience has only a very small association with student outcomes (*31*), and analysis of pre-kindergarten programs in which observed quality, not teacher or program features, predicts children's gains (*35*). Such consistent results suggest that initiatives, either as a function of regulation or professional development, that focus attention on the actual instructional and social setting of the classroom, such as mentoring, coaching, or observationally-based feedback to teachers, may be a more direct way of improving the qualities of classroom experience that make a difference for children (*36, 37*).

Despite the variation observed there is a picture of an average day in 5 [th] grade that emerges from these data. For the most part children spend time in whole group instruction or individualized seatwork focused on learning and performing basic skills (mostly in literacy and to a lesser extent math). The typical 5 [th] grader does not engage in a variety of learning experiences throughout the instructional day that includes opportunities for higher-order thinking, analysis or inference (*22, 30*); nor does he or she get to learn or work in small group or cooperative settings. The emphasis on basic skill instruction and overall high level of productivity may reflect the pressures of standards-based reform initiatives that focus on learning the correct response to structured multiple-choice test items (*38*).

Given this pattern of lack of consistency across and within grades, generally low instructional quality, and lack of association between classroom behavior and structural regulatory parameters, it is not surprising that the school effectiveness literature yields so little in the way of understanding the mechanisms underlying the gains in child achievement that do accrue as a function of school attendance (*39*).

Limitations

These findings are limited by several constraints in the study's design. First, the observations, although fairly comprehensive, did not include the entire scope of codes that might be deemed important or relevant. It is also the case that the codes themselves were somewhat limited and restrictive, such as is the case for "child engagement" that reflects only limited exposure to an activity rather than a rich and active engagement with a challenging task (e.g., *30*). Although the observations were limited to one day and did not systematically sample variation across multiple days, our sub-analyses indicate the estimates to be reliable indicators of

what occurs in a given classroom (*2*). It is also possible that the children we observed may not have been "typical" in some way, or attributes of the observed children may somehow have been related to the observations of these classrooms. However, to the extent that the NICHD SECCYD sample reflects a diverse range of family backgrounds as present in the United States, these observations reflect a typical day in school for a large number of children. It is important to note the NICHD SECCYD sample excluded children experiencing a range of potential risk conditions and therefore these descriptive results may not reflect the even wider range of educational experiences of children raised in high risk circumstances.

In summary, as was seen for the first (*1*) and third grade (*2*) classrooms of the children enrolled in the NICHD SECCYD, experiences in 5th grade are variable from classroom to classroom (and year to year) and they reflect an approach to instruction highly oriented toward teaching basic skills through whole group or individual seatwork practices. Perhaps reflecting the national emphasis on performance of academic skills on standardized tests (*33*), basic-skill activities continue to dominate the instructional landscape, even in 5th grade, when it is generally accepted that most children have acquired basic reading and math skills. Not to be overlooked is the fact that these classrooms are typically emotionally warm and positive places for the children enrolled and there appears to be a somewhat more pronounced emphasis on instruction than there was in 1st and 3rd grades.

References

1. NICHD Early Child Care Research Network [ECCRN], *Child Dev.* **71**, 960 (2002).

2. NICHD ECCRN, *Elem. Sch. J.* **105**, 305 (2005).

3. National Center for Education Statistics [NCES], *The Condition of Education 2003* (Dept. of Educ., Washington, DC, 2003).

4. NCES, *School and Staffing Survey, 1993-1994: Principals' Survey* (Dept. of Educ., Washington, DC, 1994).

5. B. K. Hamre & R. C. Pianta, *Child Dev.* **76**, 949 (2005).

6. C. Howes *et al., Dev. Psych.*, in press.

7. NICHD ECCRN, *AERJ-TLHD* (under review).

8. A. Bandura, *Self-Efficacy: The Exercise of Control* (Freeman, New York, 1994).

9. R. Crosnoe, M. K. Johnson, & G. Elder, *Soc. of Educ.* **77,** 60 (2004).

10. D. Wechsler, *Wechsler Abbreviated Scale of Intelligence* (Psychological Corp., San Antonio, 1999).

11. F. Gresham & S. Elliott, *The Social Skills Rating System* (Amer. Guidance, Circle Pines, MN, 1990).

12. T. M. Achenbach, *Manual for the Teacher's Report Form and 1991 Profile* (Univ. of Vermont, Burlington, 1991).

13. R. W. Woodcock & M. B. Johnson, *Woodcock-Johnson Psycho-Educational Battery-Revised* (Riverside, Itasca, IL, 1989).

14. NICHD ECCRN, *Amer. Educ. Res. J.* **39**, 133 (2002).

15. M. McCaslin *et al., Comprehensive School Reform: An Observational Study of 145 Teachers*, under review.

16. R. Pianta *et al., Applied Dev. Sci.* **9**, 144 (2005).

17. Appalachian Regional Advisory Committee, *Education Challenges and Technical Assistance Needs of the Appalachian Region* (CAN Corp., Alexandria, VA, 2005).

18. Educational Testing Service [ETS], http://tfp.ets.org/pub/corp/position_paper.pdf; 2004 (15 February 2006).

19. T. J. Lasley, D. Siedentop, & R. Yinger, *J. of Teacher Ed.* **57**, 13 (2006).

20. V. E. Lee & D. T. Burkham, *Inequality at the Starting Gate* (Econ. Policy Inst., Washington, DC, 2004).

21. U.S. Department of Education, *Definitions from the No Child Left Behind Act,* www.ed.gov/policy/elsec/leg/esea02/pg107.html#sec9101 (2002).

22. K. Bogner, L. Raphael, & M. Pressley, *Sci. Studies of Reading,* **6** 135 (2002).

23. 13. J. Brophy & T. Good, in *Handbook of Research on Teaching,* M. Wittrock, Ed. (MacMillan, New York, 1986), pp. 328-375.

24. D. Bryant, R. Clifford, & E. Feinberg, *Amer. Ed. Res. J.* **28**, 783 (1991).

25. L. A. Meyer, J. L. Wardrop, C. N. Hastings, & R. L. Linn, *J. of Ed. Res.* **86**, 142 (1993).

26. R. C. Pianta, K. La Paro, C. Payne, M. J. Cox, & R. Bradley, *Elem. Sch. J.* **102**, 225 (2002).

27. W. L. Sanders & J. C. Rivers, "Cumulative and residual effects of teachers on future student academic achievement" (University of Tennessee Value-Added Research and Assessment Center, Knoxville, 1996).

28. NICHD ECCRN, *Classroom Environment Predictors of Achievement Trajectories, K-5*, Manuscript under review.

29. L. Darling-Hammond, *Doing What Matters Most: Investing in Quality Teaching* (National Comm. on Teaching & America's Future, New York, 1997).

30. S. E. Dolezal, L. M. Welsh, M. Pressley, & M. M. Vincent, *Elem. Sch. J.*, **103**, 239 (2003).

31. C. T. Clotfelter, H. F. Ladd, & J. L. Vigdor, Duke University paper presented at the Economic Department, University of Virginia, Charlottesville, 2003.

32. D. D. Goldhaber & D. J. Brewer, in *Better Teachers, Better Schools*, M. Kanstoroom, C. Finn, Eds. (Fordham Foundation, New York, 1999).

33. NCES, *Overview and Inventory of State Education Reforms: 1990-2000* (Dept. of Educ., Washington, DC, 2003).

34. S. G. Rivkin, E. A. Hanushek, & J. F. Kain, *Teachers, Schools, and Academic Achievement* (Univ. of Texas at Dallas, Dallas, 2000).

35. A. Mashburn *et al., Pre-K Program Standards and Gains in Academic and Language Skills*, Manuscript under review.

36. B. K. Hamre, in *The New American Elementary School*, R. Pianta, M. Cox, & K. Snow, Eds. (Brookes, Baltimore, in press).

37. R. C. Pianta & J. P. Allen, paper presented at the Wm. T. Grant Foundation Conf. on Improving Social Settings, New York, 18 May 2006.

38. D. C. Berliner, in *Using What We Know about Teachers*, P. Hosford, Ed. (Assoc. for Supervision & Curr. Dev., Alexandria, VA, 1984)

39. M. Rutter & B. Maughan, *J. of Sch. Psych.* **40**, 451 (2002).

Table 1

Characteristics of Teachers and Classrooms Observed in 5th Grade (n=737)

	x	S.D.	Minimum	Maximum
Classrooms				
Observed # teachers	1.14	0.30	0.38	3.00
Observed # other adults	0.08	0.18	0.00	1.19
Observed # children	21.09	4.98	10.25	46.44
Observed child-teacher ratio	19.42	5.06	5.13	35.63
Teachers				
Years of public school experience	12.47	10.72	0.00	40.00
Monthly salary	4091.23	1343.54	1333.00	8778.00
Age	43.50	11.25	22.00	73.00

Table 2

Descriptions of 5th Grade Classroom Settings: Means and Mean Percent of Intervals in which Time-Sampled Codes Were Observed

	# of intervals	S.D.	Range	Mean percent of intervals
Setting				
Whole class	31.68	10.26	0 - 57.00	52.80
Large group > 5	0.87	2.98	0 - 29.25	1.45
Small group ≤ 5	4.42	5.64	0 – 35.25	7.37
Individual	23.04	9.77	0 – 54.00	38.40
Activities				
Literacy/language arts	21.97	8.64	0 – 48.00	36.62
Word-level	10.03	8.12	0 – 48.00	16.72
Comprehension	16.93	9.39	0 – 58.50	28.22
Mathematics	14.64	6.45	0 – 42.75	24.40
Computation	10.51	7.27	0 – 42.00	17.52
Concept/problem-solving	4.23	5.68	0 – 35.25	7.05
Science	6.81	6.88	0 – 37.50	11.35
Social studies	7.85	7.02	0 – 36.00	13.08
Enrichment	1.50	3.57	0 – 30.00	2.50
Technology	1.30	3.39	0 – 22.50	2.17
Free time	0.33	1.57	0 – 19.50	0.55

	# of intervals	S.D.	Range	Mean percent of intervals
Transition/management	10.25	5.77	0 – 39.00	17.08
Teacher behavior				
Attends to child	5.05	3.34	0 – 21.75	8.42
Teaches basic skills	18.99	8.79	0 – 44.25	31.65
Teaches analysis/inference	3.97	4.65	0 – 27.00	6.62
Managerial instructions	10.78	5.23	0 – 30.75	17.97
Disciplines	0.59	1.37	0 – 16.50	0.98
Positive affect	1.49	2.31	0 – 18.75	2.48
Negative affect	0.24	0.88	0 – 9.75	0.40

Table 3

Descriptions of 5th Grade Classrooms: Global Ratings

Rating	x	S.D.	Minimum	Maximum
Richness of instructional methods	3.63	1.06	1.00	6.38
Classroom overcontrol	2.15	1.03	1.00	6.50
Classroom chaos	1.40	0.54	1.00	4.50
Teacher detachment	1.83	0.70	1.00	5.38
Positive classroom climate	5.14	0.68	2.50	7.00
Negative classroom climate	1.30	0.50	1.00	4.88
Evaluative feedback	3.44	1.06	1.00	6.38
Productive use of instructional time	4.90	0.89	1.88	6.88
Teacher sensitivity	4.85	0.79	1.63	6.63

Table 4

*Descriptions of Child Behavior in 5*th *Grade: Means and Mean Percent of Intervals in which*

Time-Sampled Codes Were Observed

	x	S.D.	Minimum	Maximum	Mean percent of intervals
Engaged in learning	41.16	8.23	11.25	57.75	68.60
Highly engaged	0.10	0.54	0.00	7.50	0.17
Unproductive	18.07	8.16	1.50	48.75	30.12
Learning basic skills	34.86	9.88	0.00	58.50	58.10
Learning analysis/inference	7.68	7.72	0.00	42.75	12.80
Collaborative work	2.85	4.33	0.00	24.00	4.75
Positive/neutral with peers	12.74	8.34	0.00	41.25	21.23
Negative with peers	0.19	0.75	0.00	10.50	0.32
Positive/neutral with teacher	0.65	1.21	0.00	10.50	1.08
Negative with teacher	0.04	0.22	0.00	2.25	0.07

Table 5

Associations between Observed Quality and Teacher/School and Family/Child Attributes

	Positive social climate	Classroom instructional quality
Teacher-school characteristics		
Teacher-reported class size	-.073*	.026
Years of public school teaching experience	-.028	-.10*
Teacher's education	.007	.021
In-service related to teaching	.003	-.009
Principal involvement	.023	.020
Teacher's monthly salary	.011	.084*
Teacher's influence on school policy	.066*	.058*
Self efficacy	.181**	.126**
Family characteristics		
Maternal education	.147**	.116**
Income-to-needs 4th grade	.220**	.187***
Child characteristics		
Gender (male)	.066	.050
Externalizing 4th grade	-.170**	-.071
Social skills 4th grade	.139**	.090**
Cognitive ability 4th grade	.189**	.125**
Positive perception of school	.123**	.111**

	Positive social climate	Classroom instructional quality
Secure relationship with teacher	.111**	.102*

* $p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Table 6

Cross-Time Correlations of Classroom Quality Composites: Whole Sample (n=900) below diagonal / same school (n=485) above diagonal

	Grade 1		Grade 3		Grade 5	
	Emotional	Instructional	Emotional	Instructional	Emotional	Instructional
Grade 1						
Emotional		0.59	0.24	0.18	0.22	0.13
Instructional	0.57		0.19	0.17	0.11	0.06
Grade 3						
Emotional	0.17	0.13		0.52	0.28	0.17
Instructional	0.14	0.12	0.55		0.10	0.11
Grade 5						
Emotional	0.18	0.07	0.25	0.11		0.58
Instructional	0.11	0.05	0.13	0.09	0.56	

Table 7

Overlap in classification of stable high or low emotional or instructional quality with different levels of achievement, and income-to-needs

	Emotional Quality				Instructional Quality			
	High		Low		High		Low	
	n	%	n	%	n	%	n	%
Achievement @ 54 Months	χ^2 (2, N = 338) = 11.65**				χ^2 (2, N = 327) = 26.07***			
High	71	59.2	49	40.8	69	55.6	55	44.4
Med	45	43.3	59	56.7	42	44.7	52	55.3
Low	43	37.7	71	62.3	25	22.9	84	77.1
Income-to-Needs	χ^2 (2, N = 353) = 21.03****				χ^2 (2, N = 339) = 20.20****			
Poor	9	24.3	28	75.7	3	8.6	32	91.4
Near Poor	15	28.8	37	71.2	13	33.3	26	66.7
Not Poor	145	54.9	119	45.1	125	47.2	140	52.8

Note: * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$; **** $p < 0.0001$.

Learning, University of Virginia, 350 Old Ivy Way Suite 100, Charlottesville, VA 22903;

FAX: 434.243.5480; email: rcp4p@virginia.edu; phone: 434.243.5483.