

Standardized Classroom Observations from Pre-K to Third Grade:
A Mechanism for Improving Quality Classroom Experiences During the P-3 Years

Robert C. Pianta

University of Virginia

Curry School of Education

October, 2003

Standardized Classroom Observations from Pre-K to Third Grade:
A Mechanism for Improving Quality Classroom Experiences

During the P-3 Years

Robert C. Pianta

University of Virginia

Overview

The widespread and systematic use of standardized observation in classrooms from pre-kindergarten to third grade has the potential to address major shortcomings in educational services provided to young children. At present, there are no standardized measures for assessing pre-kindergarten/early elementary classrooms' contributions to child outcome that could be appropriately used in a system of accountability or professional development focused on *classroom* quality. Absent an appropriate tool for observational assessment, there is no way of gauging whether programs are actually offering quality experiences nor are there ways to improve program quality that focus on the actual experiences of children and teachers' behavior in classrooms.

The descriptions of classroom quality available in standardized classroom observation systems provide a standard way of measuring and noting teachers' strengths and weaknesses and evaluating whether policy initiatives or professional development activities are actually helping improve classroom interactions. This paper reviews the research on classroom observation in relation to policy and practice issues in early education, proposes a model for the use of systematic classroom observation, and describes the development of the Classroom Assessment Scoring System (CLASS). CLASS (Pianta & La Paro, 2003) describes and measures classroom quality from Pre-k to Grade 3 using a common set of dimensions and rating scales, with grade-specific examples reflecting both instructional and socioemotional aspects of the classroom environment.

Conceptual, Theoretical, and Policy Issues

The Policy/Practice/Professional Development Context for Classroom Observation

Issues being debated within several policy arenas provide support for the need for a more systematic and focused effort to document the quality of young children's experiences in early education classrooms. These involve accountability and readiness, improving program quality and access to quality educational programming for young children, and professional development.

Accountability and readiness. In the current education policy climate, there is a strong emphasis on accountability, particularly that schools and classrooms should be held more responsible for child outcomes, e.g., child achievement. However, the mechanisms for ensuring accountability rest entirely on direct assessment of children, typically starting in third grade, but increasingly initiated at younger ages, even extending into the pre-kindergarten years. This tendency to view direct assessment of children as the *only* means for ensuring accountability of classrooms and schools is quite limited, especially for young children when school readiness is the focus of assessment, policy, and programming. The technical adequacy (reliability, validity) of direct assessments for young children is widely recognized as lower than for older children, in large part, because children's competencies are somewhat unstable and situationally dependent (La Paro & Pianta, 2001). Furthermore, because the competencies of children are, in large part, dependent on the quality of their experiences in educational (and family) settings, it is logical to assess, for accountability purposes, the quality of those settings—in short to have accountability standards for classrooms (Pianta, 2003).

Programs of research establish that the kind of instruction and interactions with adults that occur in pre-kindergarten and early elementary settings have reliable and detectable effects on children's achievement and social competence (e.g., Barnett, 1995; Howes, Phillipsen, & Peisner-Feinberg, 2000; Meyer, Waldrop, Hastings, & Linn, 1993; Morrison, 1999; National Institute for Child Health and Human

Development [NICHD] Early Child Care Research Network [ECCRN], 1996; in press a; in press b; Peisner-Feinberg & Burchinal, 1997; Ripple, Gilliam, Chanana & Zigler, 1999). Although much of this work has focused on pre-k, kindergarten, and first grade settings, recent studies clearly demonstrate that characteristics of child-teacher interactions, such as feedback and warmth, produce gains in children's performance in third grade classrooms as well (Matsumura, Patthey-Chavez, Valdes, & Garnier, 2002; NICHD ECCRN, 2003). These findings strongly support the view that for young children, experiences in classrooms *matter* and that interactions between children and teachers are a primary mechanism through which classroom experiences have effects on development. Yet, there are no standardized measures that focus on child-teacher interactions for assessing classrooms' contributions to child competence and that could be used appropriately in a system of accountability focused on *classroom* quality across the P-3 years.

Access to quality that matters. It is nearly universally recognized that access to high quality educational programs is a key component of improving child outcomes. Almost every piece of state legislation that provides support for the implementation and expansion of pre-kindergarten programs for four-year-olds emphasizes that such programs must be of high quality (Bryant, Clifford, Pianta, Early, Howes, Barbarin, & Burchinal, 2002). In addition, Federal legislation such as "No Child Left Behind" emphasizes the importance of high quality instruction in the elementary grades. In fact, it would be difficult to find any policy document pertaining to the education of young children that did not emphasize the importance of high quality programming. Despite this universal recognition of the importance of quality, there is little attention to documenting whether children are attending high quality programs or ways of improving program quality that focus on the *actual interactions between children and teachers in classrooms*—the form of quality that matters for child development. Most ways of measuring and improving program quality rely on distal quality indicators such as teacher credentials or class size, which

albeit to some extent related to observed quality, are by no means strongly related or equivalent to observed quality (NICHD ECCRN, 1999; 2002; in press a). Distal quality metrics are disconnected from children's and teachers' actual classroom interactions.

Professional development linked to actual practice improvement. Few members of the educational community doubt that the development of the professional workforce in early education is of paramount importance for the delivery of quality services (Casparly, 2002). The combination of a shortage of teachers, expansion of educational programming slots for pre-kindergarten children, and an increased emphasis on instruction, particularly in literacy, have combined to make professional development (both in-service and pre-service) a linchpin in the development of quality educational programming for children below age 8. However, professional development for quality improvement typically occurs in the absence of a direct link to actual teaching behavior in classrooms, particularly for already-trained and certified teachers (Casparly, 2002). In most cases, the assessed outcome of most professional development activities is whether or not the teacher attended the activity, not whether actual practice or quality improved. Systematic observation systems provide a standard way of measuring and noting teachers' strengths and weaknesses and evaluating whether professional development activities are actually helping improve classroom interactions.

This connection among professional development, efforts to ensure equal access to quality experiences, and a system for observing quality in early learning and care environments has a long history, mainly in child care. In child care, these linkages have relied on use of the Early Childhood Environment Rating Scale (ECERS, Harms, Clifford, & Cryer, 1998) and related instruments. Throughout the country, states have relied on observations using the ECERS to track levels of and access to quality child care and have linked the ECERS assessments to professional development for child care providers. Despite this track record of using observational assessments of settings in the world of child

care, systems of professional development from pre-kindergarten to grade 3 do not directly link to observable classroom processes.

Research on Classroom Observation: What Do We Know About Children's P-3 Experiences?

During the past ten years, several large-scale observational research efforts related to schooling and young children have reported results pertaining to the quality of experiences offered to children from pre-k to third grade. Importantly, this observational work has relied on extensive research and development conducted over the course of the last several decades regarding the nature and aspects of experiences in preschool and elementary classrooms that produce gains for children (e.g., Ramey, Campbell, Burchinal, Skinner, Gardner, & Ramey, 2000) and studies on how to best measure those aspects of classroom process. This work resulted in a body of knowledge on how to define and measure quality that informed large-scale research on hundreds of classrooms from pre-k to third grade. These recent observational studies include over 240 pre-kindergarten classrooms in six states (Bryant et al., 2002), 223 kindergarten classrooms in three states (Pianta, La Paro, Payne, Cox, & Bradley, 2002), over 900 first grade classrooms in 295 school districts in 32 states (NICHD ECCRN, 2002), and over 800 third grade classrooms in more than 35 states (NICHD ECCRN, 2003). In all, there are available results from recent observational reports describing almost 2,500 early education and elementary school classrooms. As a set, these results provide the best estimate to date of the state of American education in the primary grades in terms of the actual experiences offered to children. A by-product of the rigorous training and reliability regimens necessary to conduct these objective, standardized observation studies has been a set of extensively field-tested procedures for measurement, psychometrics, and validation that are a tremendous resource for applications of observational technology at scale. In short, because of these studies, there is now available the knowledge and capacity to observe reliably in hundreds and thousands of early elementary classrooms.

The overwhelming conclusion to be drawn from these observational studies is the exceptional variability in the nature and quality of learning experiences offered to children in the early grades (see NICHD ECCRN, 2002; 2003; Pianta et al., 2002, for specific results). In the NICHD studies in kindergarten, first, and third grade, that used both global ratings and discrete time-sampled codes, the typical child received instruction in large-groups. However, in some classrooms children were never taught in a whole group, while in others this was the mode of instruction all day. Similarly, although literacy instruction was the predominant activity offered to children, in a substantial number of classrooms children were offered no literacy activities at all. Further, when rating the richness of literacy instruction in first grade classrooms, over 20 percent of classrooms were rated “poor” on the level and quality of literacy experience offered to children. This was true whether the classroom was a kindergarten, first, or third grade room (NICHD ECCRN, 2002, 2003) and was independent of the availability of materials and quality of the physical environment.

It is difficult to characterize the level of variability in these classes other than to say the entire range of codes is reflected in all the classroom samples observed. Across classrooms, nearly the entire range of scale points was observed on various global and discrete codes. In some kindergarten or first-grade classrooms, 100 percent of the activity observed in a morning-long block was organized as teacher-directed instruction. In others, none of the activities observed could be coded as this form of instruction. In some of these classrooms, children were engaged in center activities for the entire observation and in others center activities never occurred. How teachers interacted with a target child was also highly variable. In some classrooms, a typical child might interact with her teacher only as a function of being a member of the entire class during whole group instruction, while in others interaction occurred in a mix of small group and one-to-one formats. On average, there was little emphasis in these classrooms on teaching social conventions, and little evidence of affect displayed by the teacher toward a typical child.

In third grade, some classrooms were observed in which all instruction occurred with the teacher standing at the front of the room talking to the whole class while in others children spent all their time working alone at their desks (NICHD ECCRN, 2003). Also in third grade, it is still the case, even when observing for an entire day's academic activities, that many children (almost 15 percent) were not exposed any form of literacy instruction (NICHD ECCRN, 2003).

The observations in the NICHD study used a morning-long or one-day observation period, before concluding that that variation being observed was reflective of stable aspects of classrooms and not simple part of day-to-day changes in classroom activities, a final check was made on the reliability of the observations in third and first grade, by examining the associations between codes for classrooms observed more than once. Because some classrooms were attended by more than one study child, there was a sample of classrooms to examine stability of the observations across days/different children. The average cross-day correlation ranged from .71 to .91 for both third and first grade, for global ratings and for time-sampled codes (NICHD ECCRN, 2003). Thus it appears that morning-long or one-day observations are reflective of aspects of the classroom setting that remain stable across days (and different children).

In the recent National Center for Early Development and Learning (NCEDL) observational study of 240 pre-kindergarten classrooms, in which the Pre-K version of the CLASS (Pianta & La Paro, 2003) was used, similar results were reported (Bryant et al., 2002). That is, these pre-kindergarten classrooms showed exceptional variability in the nature and quality of educational programming offered to young children, and, on average, the instructional quality of the setting was fairly low. Also there was only a very small association between observed quality and policy-relevant factors such as teacher training or experience, and in general the classrooms appeared only moderately productive in the sense of providing

for active, engaging learning experiences for children (Pianta, Howes, Burchinal, Bryant, Clifford, Early, & Barbarin, 2003).

Despite this variability, a picture emerges of the general, or “average” PreK-3 early education classroom: instruction delivered in a whole-group setting, a positive social environment, and low levels of child productivity and engagement in academic activities. These environments can be characterized as socially positive but instructionally passive: children listen and watch; much time is spent on routines or management of materials; and children have little direct contact with teachers in instructional interactions. This actually seems to be increasingly the case as children move from first to third grade—the classroom becomes even more passive, in the sense that in third grade over 90% of observed occasions on a typical day involved whole-class instruction or individual seatwork (NICHD ECCRN, 2003). Despite being generally well-organized and busy places, classrooms appear low on “intentionality,” a term that refers to directed, designed interactions between children and teachers in which teachers purposefully challenge, scaffold, and extend children’s skills. Another major finding across all these studies and grade-levels is that factors used to regulate classroom quality such as teacher education or class size, as well as the use of any or a specific curriculum, bear little or no relation to observed quality or to child outcomes (Pianta, 2003).

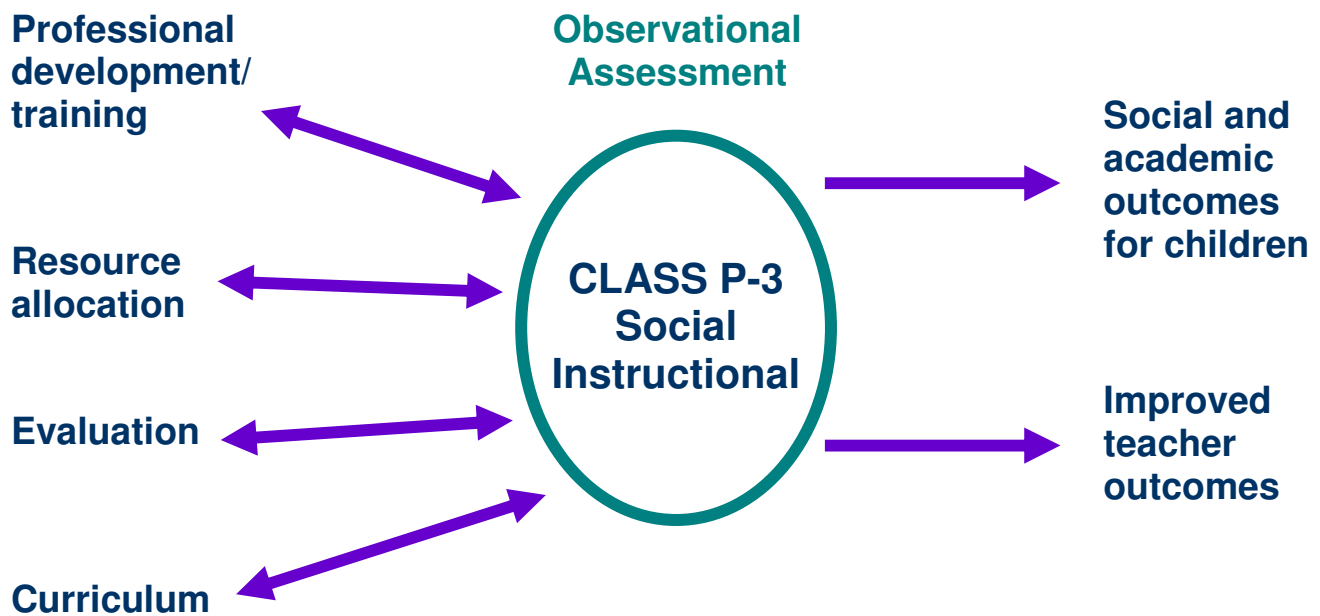
Several implications follow from this work. The finding of high between-classroom variability, reported now in several papers from the NICHD Study of Early Child Care and Youth Development (SECCYD) (NICHD ECCRN, 2002; Pianta et al., 2002) and a range of other somewhat smaller investigations (e.g., Bryant, Clifford, & Feinberg, 1991; Meyer et al., 1993) suggests little agreement on what constitutes an appropriate instructional program for elementary school children: how to find the correct balance of instructional and social activities, the best way to distribute time across a range of content areas and activities, and how to deliver information and instruction to young children (e.g.,

Bredekamp & Copple, 1997; File & Gullo, 2002; NEGP, 1998). This variability is present across schools and across grades and classrooms within the same school. These results raise doubts about access to quality experiences, equity of those experiences for various segments of the population, and the apparent irrelevance of regulable features of classroom settings.

Shaping the Future: Mechanisms for Improving Access to Quality Learning Experiences

Systematic observational instruments used in P-3 classrooms are designed to build on current research and address problems in the delivery of educational services to young children. Observation is a mechanism with the potential to link educational policies (e.g., class size, financial resources), or professional development (e.g., in-service or pre-service training) with the *actual* experiences of children in classrooms. Figure 1 (below) depicts a model in which the systematic use of classroom observation

Figure 1: Linking policy and professional development with teacher and child outcomes through standardized observations of classroom processes.



systems can function as a tool for describing how curriculum is implemented in classrooms. These descriptions, when linked to policy and training questions, can contribute to systematic improvements in how resources contribute to improvements in classroom quality and child outcomes.

In Figure 1, educational policy-relevant factors are linked with outcomes for children and for teachers as a function of the social and instructional qualities of the classroom environment, in this case assessed by the CLASS (Pianta & La Paro, 2003). The model hypothesizes that decisions about policy factors such as class size, curriculum, professional development, evaluation of students, or teachers affect children's performance and teachers' outcomes (such as leaving or staying in the profession), as a function of changes these policies make in the classroom environment. In this model, the use of a standardized metric for classroom experiences allows for evaluating on a systematic basis, the extent to which manipulations and variations in policy result in the expected changes in classroom process and child/teacher outcomes. Interestingly, recent reviews of school effects suggest the link between policy and outcome is quite small *likely due to the very high levels of variability in the mediating factor of classroom processes* (Rutter & Maughan, 2002).

As an example of policy→process→outcome linkage, in child care settings increments in training and professionalism of care providers as well as reduced child-teacher ratio are clearly related to more positive child outcomes as a function of improvements in child care quality (NICHD ECCRN, 1999). However, the effects are small. In elementary schools, it has been demonstrated that reductions in class size, particularly in British schools, have been linked to more positive academic and social performance as a consequence of improved instruction, time on task, and opportunities for learning in the classroom setting (Blatchford et al., 2002a; 2002b). Several analyses of state-wide assessments of children's achievement conclude that positive increments in outcome are a function of "teacher quality" or

classroom-level mechanisms (Hanuschek, 1999) which in survey studies go unmeasured but which are likely to include indicators of classroom processes and instructional quality.

What is critical from a policy change standpoint is the way that measures of classroom process such as that available with the CLASS, can over time assist in producing gains for children and teachers in a systematic way. For example, by using observation to assess classroom processes, policy makers and trainers can gauge the need for alterations in training, curriculum implementation, or resource allocation that, in turn, could produce better classroom quality. This “feedback loop” between classroom quality and policy is currently underutilized and underexamined largely as a consequence of the lack of standardized measurement of classroom processes being implemented in schools. In short, states, trainers, and school divisions have operated without tools to link policy and training to children’s and teachers’ experiences in classrooms and hence have little idea how policies or training may or may not work and little information regarding if they work to produce change in classrooms.

CLASS has been developed to fill the need for such an instrument from pre-k through the early elementary grades. In child care settings, the Early Childhood Environment Rating Scale (Harms & Clifford, 1983; Harms et al., 1998) and the Infant and Toddler Environment Rating Scale (Harms, Cryer, & Clifford, 1987), have functioned as assessment tools linking policy and training to setting quality and child outcomes for several years now. States have adopted these instruments as part of professional development quality indicators and as part of program and classroom certification efforts. The widespread and successful introduction of these instruments into state and regional child care evaluation and training efforts has been shown to be related to improvements in quality and associated child outcomes. Thus there exists a precedent in the education and care of young children for assessment of the environment, and standards of quality associated with that assessment, to be a mechanism for linking policy and training to outcomes and for systematic improvement. As noted, no such procedure exists for

early elementary classrooms, despite a nearly decade-long focus on school readiness and ready schools. The ECERS, by and large, is not as well suited for use in elementary school settings as it is in child care, because of the increased focus in the elementary grades on instruction and content, and the reliance of the ECERS on aspects of the physical environment and safety for measuring quality. Although an exceptionally valuable tool for assessing child care environments, the ECERS was not developed for and may not be appropriate for, use in observing quality in kindergarten and the elementary grades.

Assessing the Quality of Classrooms as Learning Environments: Development of the CLASS

Conceptualizing classroom quality. Work on defining and operationalizing quality in early education and elementary classroom environments ranges from summaries and recommendations of what constitutes developmentally appropriate practice (Bredekamp & Copple, 1997) to empirical research on constructs such as “feedback” (Matsumara et al., 2002; Meyer et al., 1993; Morrison, 1999). The term quality has been used for many years in the literature on child care and early childhood education (see Bredekamp & Copple, 1997; NICHD ECCRN, 1996; 1999). This literature is a rich source of information on measurement of the quality of children’s experiences in classroom-like settings. The child care literature also provides information on how quality, in turn, relates to a range of teacher, family, and program correlates as well as child outcomes (Bredekamp & Copple, 1997; NICHD ECCRN 1996; 1999).

A wide-ranging debate concerns how to define, measure, and improve those aspects of classroom environments that affect children’s learning. Ranging from recommendations about balancing the instructional/academic and emotional/social aspects of the classroom (Bredekamp & Copple, 1997; Maxwell, McWilliams, Hemmeter, Ault, & Schuster, 2001) to empirical research on constructs such as “feedback” (Matsumura et al., 2002; Meyer et al., 1993), there is no shortage of discourse and very little large-scale research on this issue as it is enacted in real classrooms. Although the term “quality” has been used for many years in the literature on child care and early childhood education (see Bredekamp &

Copple, 1997; NICHD ECCRN, 1996), it is a global construct that is not often or easily applied in research on elementary school classrooms, in which the focus is often on specific teaching behaviors (e.g., Meyer et al., 1993). For example, in classroom-focused research, researchers tend to focus on teachers' uses of curriculum and instructional strategies (e.g.; Blachman, Tangel, Ball, Black, & McGraw, 1999; Goffin & Day, 1994) or on specific teacher behaviors such as praise or modeling (e.g., Brophy & Good, 1986).

Berliner's work (1987) draws the distinction between good teaching and effective teaching in that although good teaching requires the teacher to understand children's development in general terms; effective teaching requires an understanding of the expected outcome of instruction or the way development changes as a function of instruction. This distinction implies a role for the "good" teacher that extends beyond an understanding the child to enacted intentionality in interactions with the child. High expectations, skills of management and planning, a learning-orientation in the classroom, engaging children's attention with appropriate activities, and using effective, evaluative feedback and questioning are all part of the intentionality of good teaching (Berliner, 1984). For example, Bogner, Rapael, and Presley (2002) examined how teachers motivate students during literacy instruction in seven first-grade classrooms and found that literacy engagement was higher in classrooms in which teachers used scaffolding, encouraged autonomy and choice, had a caring manner, interacted positively with students one-to-one, made deep and personal connections to students, encouraged creativity, and generally set a positive tone in the classroom. Similarly, Dolezal, Welsh, Pressley, and Vincent (2003) found that teachers in highly engaging third grade classrooms provided tasks that were cognitively challenging, presented probing questions, and provided a positive classroom atmosphere. Thus teachers' active, intentionally instructive behaviors as well as socially warm and responsive approach appear closely linked

to children's functioning in elementary school. Accordingly, quality is conceptualized in terms of specific teacher-child interactions that contribute to children's learning and development.

This conceptualization of quality in terms of teacher-child interactions raises a key distinction in classroom observation assessments, such as the CLASS and the ECERS, as well as other rating-based tools such as the Early Language and Literacy Classroom Observation (Smith & Dickinson, 2002). The CLASS scales are based entirely on *interactions* of teachers and children in the classroom; scoring for any scale is not determined by the presence of materials, the physical environment or safety, or the adoption of a specific curriculum. The CLASS assesses the quality of teachers' implementation and use of a curriculum, the quality of her social and instructional interactions with children, and the intentionality and productivity evident in the classroom setting. This distinction between *observed interactions* and physical materials or reported use of curriculum is important, because in most early elementary settings, materials and curriculum are usually prevalent and fairly well organized. And so unlike in child care settings that are typically quite variable on this dimension, early elementary school settings can be viewed as of higher quality because of the presence of materials or curricula when the quality of *use and implementation* of curriculum and materials is what determines the value of the classroom for developmental progress (Greenberg, Domitrovich, & Baumgardner, 2001; Morrison & Connor, 2002; Rutter & Maughan, 2002; Pianta, 2003). In the CLASS, the focus is on what teachers *do* with the materials they have and in the interactions they have with children.

A second distinction designed into the CLASS that is particularly relevant for P-3 policies and improvements is that it defines quality across these grades using the same set of global dimensions (e.g. teacher sensitivity) while operationalizing them with concrete, grade-specific anchor points. By using common global dimensions across grades, the CLASS focuses attention on classroom quality in a

consistent manner across the first four years of schooling, thus potentially contributing to increased consistency and coherence in children's experiences over time from ages four through eight.

Measuring classroom quality. Observations of quality in classroom-like settings with young children most often involve global ratings of the environment on a variety of articulated dimensions that are believed to index quality. Standardized procedures using global ratings to evaluate the classroom environment include the Early Childhood Environment Rating Scale (ECERS, Harms et al., 1998), Stipek's (1996) Early Childhood Classroom Observation Measure (ECCOM), and the Observational Record of the Caregiving Environment (ORCE) and the Classroom Observation System (COS-1; COS-3) developed for the large-scale NICHD Study of Early Child Care (see NICHD ECCRN, 1996; 2002; 2003).

The ECERS is a widely accepted observational tool and has been used reliably in many large and small-scale studies to provide a measure of global quality (Love, Meckstroth, & Sprachman, 1997). The ECERS assesses routine care needs, furnishings and display, activities and experiences related to motor, language, cognitive, and social development, and adult provisions using a seven-point Likert-type scale. Statements are provided to coders that describe a specific construct or "item" such as "routine care needs" and also describe rating levels for each item. Observers watch the classroom and can interview teachers to provide additional information and then make a rating for that item/construct. A score of 5 on an item generally indicates meeting developmentally appropriate criteria in the adequate materials and supervision are provided for the activity (Harms & Clifford, 1983). Item ratings are summed to create an overall global quality total score for each classroom. The ECERS addresses primarily structural aspects of children's learning environments including provision of activities as well as materials and space. Most investigations demonstrate that it is actually the process-items on the ECERS that consistently relate to children's gains in skills.

The Observational Record of the Caregiving Environment (NICHD ECCRN, 1996) was developed specifically to assess the quality of caregiver-child interaction experienced by an individual child. The ORCE employs both an interval-level coding system for indexing the occurrence of discrete behaviors (e.g., caregivers talking to children) and global ratings of the environment and persons in the environment (e.g., teacher/caregiver sensitivity). The ORCE measures environmental factors that are generally considered to constitute quality: care that is attentive and appropriately responsive, expresses positive affect and affection, is not excessively restrictive or intrusive, and offers activities believed to promote children's cognitive and social development.

The ORCE was further developed by the NICHD Early Child Care Research Network into the Classroom Observation System (COS; Pianta et al., 2002), with versions developed and validated for kindergarten, first grade, and third grade elementary school classrooms. The COS-K, COS-1, and COS-3 follow the format of the ORCE by using time-sampling and qualitative ratings on a seven-point scale. Importantly, the COS adds to the ORCE the dimension of instructional quality and therefore assesses the classroom environment in terms of literacy instruction, richness of instructional methods, productive use of time, and instructional interactions between the teacher and the children. Another feature of the COS, as a consequence of its development in the 10-site NICHD Study of Early Child Care, is its widespread applicability across exceptionally diverse classroom settings in 30 states and more than 700 schools, as well as the development of a set of reliability and training procedures to certify observers in diverse settings to use the instrument reliably (NICHD ECCRN, 2002; Pianta et al., 2002).

Results of several large-scale studies using these rating systems identify a set of classroom process variables that are consistently linked with child outcomes. These process variables can be observed reliably across a variety of classroom settings, serve as a starting point for consensus on the constructs

involved in an operational definition of classroom quality (Love et al., 1997), and form the basis for the specific scales used in the CLASS.

The Classroom Assessment Scoring System. The CLASS was developed based on the scales used in large-scale classroom observation studies in the NICHD Study of Early Child Care (NICHD ECCRN, 2002; Pianta et al., 2002) and the National Center for Early Development and Learning (NCEDL) Multi-State Pre-K Study (Bryant et al., 2002). The scales used in the CLASS were derived from the review of constructs assessed in classroom observation instruments used in child care and elementary school research (see above for a summary). The CLASS was designed to create a common metric or vocabulary that could be used to describe various aspects of quality *across the early grades* and to have the same constructs applied to different grades. As is evident from the above summary, a common set of constructs or scales can be identified across studies and instruments, each of which can be operationalized in a pre-kindergarten to third grade classroom settings.

At the broadest level, classroom process can be described in terms of the social/emotional climate of the classroom, the nature and quality of teacher-child interactions, and the quality of instruction. The CLASS assesses 14 constructs within these broad areas, including: (a) positive emotional climate, (b) negative emotional climate, (c) teacher sensitivity, (d) overcontrol, (e) behavior management, (f) productivity, (g) concept development, (h) instructional learning formats, (i) roteness, and (j) quality of feedback; as well as interactions between teachers and children focused on (k) literacy, (l) mathematics, (m) science, and (n) social studies. The manuals provide a detailed description of each of these constructs with specific examples and definitions tailored to each grade level. In the CLASS approach, the dimensions used to define classroom quality and on which classrooms are assessed (those 14 noted above) are *common across the P-3 grades* at the same time they ways these dimensions are manifest and demonstrated are specific to each grade. In this way, the CLASS approach provides a common metric and

language for discussion of quality *across grades* (thereby addressing problems with grade-to-grade transition and the need for coherence) at the same time providing and developmentally-sensitive metric for each grade level. The following discussion elaborates on the dimensions of classroom experience assessed by the CLASS P-3.

Socioemotional climate and quality of teacher-child interactions. The ability to maintain caring and supportive relationships with children is crucial for all teachers of young children (Pianta, 1999). Sensitive teachers and teachers that create a positive climate in their classrooms tend to more familiar with the academic needs of individual children in their classroom (Helmke & Schrader, 1988). Sensitive teacher-child interactions form the basis for the development of these supportive relationships (Kontos & Wilcox-Herzog, 1997). Pianta (1999) describes the teacher-child relationship as a regulatory system that contributes to children's social and academic competencies in school. The constructs included under emotional climate on the CLASS are: *positive climate*, which examines the enthusiasm, enjoyment and emotional connection that teacher has with the children as well as the nature of peer interactions; *negative climate*, includes evidence of anger, hostility or aggression exhibited by the teacher and/or children in the classroom; and *sensitivity*, how responsive the teacher is to children academic and emotional needs and the degree to which the teacher can be viewed as a secure base for children. These constructs collectively, and separately, predict to 1) children performance on standardized tests of literacy skills in pre-k, and grade 1 (NICHD ECCRN, 2003; Pianta, 2003); lower levels of mother-reports internalizing behaviors in kindergarten and first grade (NICHD ECCRN, in press b), and to children's engagement in the classroom across all grade levels (Bryant et al., 2002; NICHD ECCRN 2002;2003). Importantly, these studies demonstrate that these aspects of classroom experience *uniquely predict* child outcomes, adjusting for selection effects and prior child functioning.

Effectiveness is highly evident in teachers with exemplary classroom management (Morrow, Tracey, Woo, & Pressley, 1999). The CLASS assesses management with three scales: *over-control*, the flexibility that teachers display related to children's interests and classroom schedules and the degree to which autonomous behavior in children is fostered; *behavior management*, how well teachers monitor, prevent and redirect behavior; and *productivity*, how well the classroom runs with respect to routines, how well children understand the routine and the degree to which teachers provide activities and directions so that maximum time can be spent in productive learning activities. Like for sensitivity above, ratings of productivity predict uniquely to children's performance on academic achievement tests in pre-k and elementary grades. Behavior management is a significant predictor of child engagement and peer competence while over-control is related to higher levels of internalizing behavior (Bryant et al., 2002; NICHD ECCRN 2002; 2003; Pianta, 2003).

Instructional support. With respect to instructional support, highly skilled teachers monitor their students' performance and provide additional explanation and ideas (Meyer et al., 1993). Thus, the teacher provides feedback to students through scaffolding and support (Yates & Yates, 1990). The CLASS has four scales that focus on instructional support provided in the classroom: *concept development*, the degree to which teachers promote higher order thinking and problem solving, going beyond fact and recall activities for children; *instructional learning formats*, how teacher engage children in activities and facilitate activities so that learning opportunities are maximized; *roteness*, how much recitation and drill occurs in the classroom; and *quality of feedback*, how teachers extend children's learning through their responses and participation in activities. These constructs together, and quality of feedback in particular, form an index of the instructional value of the classroom that predicts to academic functioning of the child in literacy and general knowledge as cited above, with indications that this association is due to increased child engagement in the activity (NICHD ECCRN, 2003). These scales

tend to be less strongly associated with social functioning than do the assessments of emotional quality (NICHD ECCRN, 2003).

Of note is that the multi-grade CLASS rating system also includes scales that correspond to academic content in four areas: *Literacy*, *Mathematics*, *Social Studies*, and *Science*. These scales were developed to capture the increasing amount of time and instruction devoted to these skill domains as the child moves from pre-k through to Grade 3. The literacy and mathematics scales have been field-tested in the NCEDL and NICHD studies and correlate, as expected, with children's performance in those areas (NICHD ECCRN, 2003). The scales for Social Studies and Science are viewed as descriptive at the present time, as there are no validity data for those indicators.

In sum, the entire CLASS system is composed of ten quality process scales as well as four domain-specific instructional scales. A developmentally or grade-level-appropriate description for each scale is provided in the manual, with specific anchor points described.

Summary and Future Directions

Systematic observations of pre-k to Grade 3 classrooms provide a mechanism for linking important policy and program development questions to child and teacher outcomes. In this way, observations of classrooms fill a critical niche in the development of a system capable of making measurable and observable changes in classroom experiences that produce developmental gains for children. Also, large-scale observation studies have shown that stand-alone training materials and coding guides can be developed and used successfully to ensure reliable coding of these constructs using rating scale systems.

Realizing the potential of observation of classroom processes for policy and professional development requires: 1) a measurement tool for conducting such observations, 2) a set of standardized procedures for ensuring quality control in training and reliable use of such a tool, 3) materials by which

educators can use this tool as a means of program development and evaluation, and eventually professional development and training, and 4) dissemination of these materials to appropriate consumers.

Existing infrastructure clearly supports development and widespread use of observation in P-3 classrooms and forms the basis for the CLASS P-3 system of observation tools. These tools potentially fill an important void in the scaling up of pre-kindergarten programs, addressing issues of access, equity, and coherence in children's schooling experiences in the early elementary grades, and in creating systems of quality improvement in which professional development is tied directly to individual teachers' practices.

References

- Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *Future of Children, 5*.
- Berliner, D. C. (1987). Simple views of effective teaching and a simple theory of classroom instruction. In D. C. Berliner and B. V. Rosenshine (Eds.), *Talks to teachers: A festschrift for N. L. Gage* (pp. 93-110). New York: Random House.
- Berliner, D. C. (1984, February). Remarks to the Governor's Task Force on Teacher Education, University of Arizona, Tucson.
- Blachman, B. A., Tangel, D. M., Ball, E., Black, R., & McGraw, C. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing: An Interdisciplinary Journal, 11*, 239-273.
- Blatchford, P., Goldstein, H., Maratin, C., & Browne, W. (2002). A study of class size effects in English school reception year classes. *British Educational Research Journal, 28*(2), 171-187.
- Blatchford, P., Moriarty, V., Edmonds, S., & Martin, C. (2002). Relationships between class size and teaching: A multimethod analysis of English infant schools. *American Educational Research Journal, 39*(1), 101-132.
- Bogner, K., Raphael, L., & Pressley, M. (2002). How grade 1 teachers motivate literate activity by their students. *Scientific Studies of Reading, 6*(2), 135-165.
- Bredekamp, S., & Copple, C. (Eds.). (1997). *Developmentally appropriate practice in early childhood programs* (Rev. ed). Washington, DC: National Association for the Education of Young Children.

- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. L. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-275). Indianapolis: Macmillan Publishing USA.
- Bryant, D., Clifford, R., Early, D., Pianta, R., Howes, C., Barbarin, O., & Burchinal, M. (2002, November). *Findings from the NCEDE Multi-State Pre-Kindergarten Study*. Paper presented at the annual meeting of the National Association for the Education of Young Children, New York, NY.
- Bryant, D., Clifford, R., & Feinberg, E. (1991). Best practices for beginners: Developmental appropriateness in kindergarten. *American Educational Research Journal*, 28 (4), 783-803.
- Caspary, K. (2002). California's pioneering training and retention initiatives for early childhood educators. Berkeley, CA: PACE.
- Dolezal, S., Welsh, L., Pressley, M., & Vincent, M. (2003). How nine third-grade teachers motivate student academic engagement. *The Elementary School Journal*, 103(3), 239-269.
- File, N., & Gullo, D. F. (2002). A comparison of early childhood and elementary education students' beliefs about primary classroom teaching practices. *Early Childhood Research Quarterly*, 17(1), 126-137.
- Goffin, S. G., & Day, D. E. (1994). *New perspective in early childhood teacher education: Bringing practitioners into the debate*. New York: Teachers College Press, Columbia University.
- Greenberg, M. T., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention and Treatment*, Vol. 4.
- Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size. *Educational Evaluation and Policy Analysis*, 21, 143-163.

- Harms, T., & Clifford, R. M. (1983). *Early childhood environment rating scale*. New York: Teachers College Press.
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale* (Rev. ed). New York: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. M. (1987). *Infant/Toddlers Environment Rating Scale* (ITERS). New York: Teachers' College Press.
- Helmke, A., & Schrader, F. W. (1988). Successful student practice during seatwork: Efficient management and active supervision not enough.
- Howes, C., Phillipsen, L., & Peisner-Feinberg, C. (2000). The consistency and predictability of teacher-child relationships during the transition to kindergarten. *Journal of School Psychology, 38*(2), 113-132.
- Kontos, S., & Wilcox-Herzog, A. (1997). Teachers' interactions with children: Why are they so important? Research in review.
- La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research, 70*(4), 443-484.
- Love, J.M., Meckstroth, A., & Sprachman, S. (1997). *Measuring the quality of program environments in Head Start and other early childhood programs: A review and recommendations for future research* (Contract RN 94094001). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Matsumura, L. C., Patthey-Chavez, G.G., Valdes, R. & Garnier, H. (2002). Teacher feedback, writing assignment quality, and third-grade students' revision in higher and lower achieving schools. *The Elementary School Journal, 103*, 3-25.

- Maxwell, K. L., McWilliams, R. A., Hemmeter, M. L, Ault, M. J., & Schuster, J. W. (2001). Predictors of developmentally appropriate classroom practices in kindergarten through third grade. *Early Childhood Research Quarterly, 16*, 431-452.
- Meyer, L.A., Wardrop, J.L., Hastings, C.N., & Linn, R.L. (1993). Effects of ability and settings on kindergarteners' reading performance. *Journal of Educational Research, 86*(3), 142-160.
- Morrison, F. (1999, August). Improving literacy in America: The role of school transition. Symposium conducted at the meeting of the American Psychological Association, Boston, MA.
- Morrison, F. J., & Connor, C. M. (2002). Understanding schooling effects on early literacy: A working research strategy. *Journal of School Psychology, 40*(6), 493-500.
- Morrow, L. M., Tracey, D. H., Woo, D. G., & Pressley, M. (1999). Characteristics of Exemplary First-Grade Literacy Instruction.
- National Education Goals Panel. (1995). *National education goals report executive summary: Improving education through family-school-community partnerships*. Washington, DC: National Education Goals Panel.
- National Education Goals Panel. (1998). *Ready schools*. Washington, DC: National Educational Goals Panel.
- National Institute of Child Health and Human Development, Early Child Care Research Network (1996). Characteristics of infant child care: Factors contributing to positive caregiving. *Early Childhood Research Quarterly, 11*(3), 269-306.
- National Institute of Child Health and Human Development, Early Child Care Research Network (1999). Child outcomes when child care center classes meet recommended standards for quality. *American Journal of Public Health, 89*, 1072-1077.
- National Institute of Child Health and Human Development, Early Child Care Research

Network (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behaviors. *The Elementary School Journal*, 102(5), 367-387.

National Institute of Child Health and Human Development, Early Child Care Research

Network (in press a). Social functioning in first grade: Prediction from home, child care and concurrent school experience. *Child Development*.

National Institute of Child Health and Human Development, Early Child Care Research

Network (in press b). Class size in first grade: Relations with classroom process and child outcomes. *Educational Evaluation and Policy Analysis*.

National Institute of Child Health and Human Development, Early Child Care Research

Network (2003). A day in third grade: Observational descriptions of third grade classrooms and associations with teacher characteristics. Manuscript in preparation.

Peisner-Feinberg, E.S., & Burchinal, M.R. (1997). Relations between preschool children's child-care experiences and concurrent development: The Cost, Quality, and Outcomes Study. *Merrill-Palmer Quarterly*, 43(3), 451-477.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.

Pianta, R. C. (2003, March). *Professional development and observations of classroom process*. Paper presented at the SEED Symposium on Early Childhood Professional Development, Washington, DC.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2003, April). *Preschool teachers and classrooms: Results from the National Center for Early Development and Learning six-state study*. Paper presented at the annual meeting of the Society for Research in Child Development, Tampa, FL.

- Pianta, R. C., & La Paro, K. M. (2003). *Classroom Assessment Scoring System: Guide and Training Manual*. University of Virginia, Charlottesville. Unpublished manuscript.
- Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal, 102*(3), 225-238.
- Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science, 4*, 2-14.
- Ripple, C. H., Gilliam, W. S., Chanana, N., & Zigler, E. (1999). Will fifty cooks spoil the broth? The debate over entrusting Head Start to the states. *American Psychologist, 54*, 327-343.
- Rutter, M., & Maughan, B. (2002). School effectiveness findings, 1979-2002. *Journal of School Psychology, 40*(6), 451-475.
- Smith, M. & Dickinson, D. (2002). *Early language and literacy classroom observation*. Baltimore, MD: Brookes Publishing.
- Stipek, D. J. (1996). Motivation and instruction. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 85-113). New York: Macmillan.
- Yates, G. C. R., & Yates, S. M. (1990). Teacher effectiveness research: Towards describing user-friendly classroom instruction. *Educational Psychology, 10*(3), 225-238.