6-2013

Teaching through Interactions: Testing a Developmental Framework of Teacher Effectiveness in over 4,000 Classrooms

Bridget K. Hamre
University of Virginia

Robert C. Pianta
University of Virginia

Jason T. Downer
University of Virginia

Jamie DeCoster
University of Virginia

Andrew J. Mashburn
Portland State University, mashburn@pdx.edu

Citation Details

Teaching through Interactions: Testing a Developmental Framework of Teacher Effectiveness in over 4,000 Classrooms Bridget K. Hamre, Robert C. Pianta, Jason T. Downer, Jamie DeCoster, Andrew J. Mashburn, Stephanie M. Jones, Joshua L. Brown, Elise Cappella, Marc Atkins, Susan E. Rivers, Marc A. Brackett, and Aki Hamagami The Elementary School Journal , Vol. 113, No. 4 (June 2013) , pp. 461-487

TEACHING THROUGH INTERACTIONS

Testing a Developmental Framework of Teacher Effectiveness in over 4,000 Classrooms

ABSTRACT

Validating frameworks for understanding classroom processes that contribute to student learning and development is important to advance the scientific study of teaching. This article presents one such framework, Teaching through Interactions, which posits that teacher-student interactions are a central driver for student learning and organizes teacher-student interactions into three major domains. Results provide evidence that across 4,341 preschool to elementary classrooms (1) teacher-student classroom interactions comprise distinct emotional, organizational, and instructional domains; (2) the three-domain latent structure is a better fit to observational data than alternative one- and two-domain models of teacher-student classroom interactions; and (3) the three-domain structure is the best-fitting model across multiple data sets.

Bridget K. Hamre
Robert C. Pianta
Jason T. Downer
Jamie DeCoster
UNIVERSITY OF VIRGINIA

Andrew J. Mashburn
PORTLAND STATE UNIVERSITY

Stephanie M. Jones
HARVARD UNIVERSITY

Joshua L. Brown
FORDHAM UNIVERSITY

Elise Cappella
NEW YORK UNIVERSITY

Marc Atkins
UNIVERSITY OF ILLINOIS, CHICAGO

Susan E. Rivers
Marc A. Brackett
YALE UNIVERSITY

Aki Hamagami
UNIVERSITY OF VIRGINIA

THE ELEMENTARY SCHOOL JOURNAL VOLUME 113, NUMBER 4

THREE converging forces dominate discussion about early childhood and elementary education in the United States, which collectively reflect a new era of accountability: (1) attention to the quality and productivity of America's classrooms and the central role of teachers, (2) the critical role that these early years play in closing the achievement gap, and (3) innovation in incentives and preparation structures that are linked to assessments of teachers' performance in the classroom. This article's focus is situated at the intersection of these forces. First, we argue that teachers' performance and effectiveness is in large part a function of their behavior in classrooms as they interact with students and implement curricula. Second, we present a conceptual model that organizes teacher-student interactions into three broad latent domains reflecting emotional, organizational, and instructional features of interactions. Finally, we empirically test the fit of the three-domain model across the largest sample of standardized observations in preschool and elementary classrooms available to date. Specifically, we test the degree to which the Teaching through Interactions model (Hamre & Pianta, 2007) is consistent with observational data collected in a large, diverse sample of 4,341 classrooms across the United States.

Numerous studies relying on sophisticated multilevel analyses of large-scale student achievement test outcomes indicate that a significant portion of variance in student learning is explained at the classroom level (Hanushek, 2002) and that deflections in the trajectory of student learning across years can be attributed to their experiences in specific classrooms (e.g., Sanders & Rivers, 1996). These studies support the conclusion that classroom experiences matter but fall short on two counts: (1) identifying specific processes that lead to student learning and positive social adjustment and (2) anchoring classroom effects in verifiably observable indicators (Pianta & Hamre, 2009). This identification and anchoring of specific classroom practice is critical to advancing the systematic and effective training of teachers and improving their performance in classrooms. More specifically, neither Hanushek's (2002) definition of teacher quality, "Good teachers are ones who get large gains in student achievement for their classes; bad teachers are just the opposite" (p. 3), nor the "value-added" paradigm for teacher evaluation (Sanders & Rivers, 1996) provide guidance for the systematic development of evidenced-based ways to improve classroom teaching or teacher preparation (Cochran-Smith & Zeichner, 2005). The critical unanswered question in the value-added definition of teacher effectiveness involves the mechanisms through which classrooms exert their influence on students' development and how such effects can be reliably produced and maximized. That is, if a teacher does or does not produce expected levels of value-added performance, what did the teacher do in the classroom that led to these outcomes?

These questions have come to the forefront of educational policy and research in recent years and provided the impetus for the largest study of teaching in elementary and secondary schools to date—the Measures of Effective Teaching (MET) study (Kane & Staiger, 2012). Results from the first year of the MET study provide compelling evidence that effective teaching can be reliably observed and that these observations are associated in meaningful ways with students' perceptions of teachers and with gains on standardized achievement tests. However, the results of this study also point out a need for more clarity around the components

of teaching that produce outcomes and a better understanding of the ways in which these components are organized in typical classrooms. This need for clearer articulation of theoretical frameworks for understanding teaching was suggest by Douglas (2009), who wrote, "Our knowledge of classroom instruction will be well served by studies that build on a common theoretical framework" (p. 519).

The present study tests one such framework of effective teaching. Effective teaching is sometimes described quite broadly to include dimensions such as teacher knowledge, teacher practices, teacher beliefs, student beliefs, student practices, and student knowledge (Bell, Gitomer, McCaffrey, Hamre, & Pianta, 2011). The operational definition of effective teaching used for this article is more circumscribed and focuses exclusively on interactions between teachers and students in the classroom. Although we recognize that there are many things that teachers do that may make them effective (e.g., provide assignments and homework, assess their students, and collaborate with parents and other teachers), we focus exclusively on the nature and quality of their interactions with students for several reasons. Developmental theory and research provides strong support for the idea that it is the daily interactions that children and adolescents have with adults and peers that drive learning and development (Bronfenbrenner & Morris, 1998). On a more practical note, current discussions of teacher accountability suggest the need to attend to aspects of teachers' jobs that can be reliably observed and assessed. We argue that the daily interactions teachers have with students are among the most feasible to include in this type of system. This orientation does not diminish the importance of other aspects of teachers' work but rather places an emphasis on the importance of interactions as a major part of their daily lives in the classroom.

This approach to conceptualizing teacher-student interactions (Hamre & Pianta, 2007) draws heavily from earlier theoretical and empirical work in the educational and psychological literatures (e.g., Brophy, 1999; Eccles & Roeser, 1999; Pressley et al., 2003) to describe one theory of classroom practice and the salience of interactions with adults for promoting developmental gains. This model proposes a *multilevel latent structure* for organizing teacher-student interactions; three broad domains of classroom interactions involving teachers and students are hypothesized to be important in promoting student learning and social development—Emotional Support, Classroom Organization, and Instructional Support. Each of these domains is quite broad and could include a wide array of interactions. The observational measure developed to assess these interactions, the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008),[1] describes in detail several specific dimensions of teacher-child interactions that exist within each broad domain (see Fig. 1). Table 1 provides a description of each dimension. For example, the domain of Emotional Support consists of four dimensions—Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. Each of these dimensions, in turn, is described by explicit indicators of that dimension. For example, Positive Climate is indicated by the presence of relationships, positive affect, positive communications, and respect. Finally, each indicator is further operationalized in specific behavioral, observable descriptions of classroom interactions, either be-

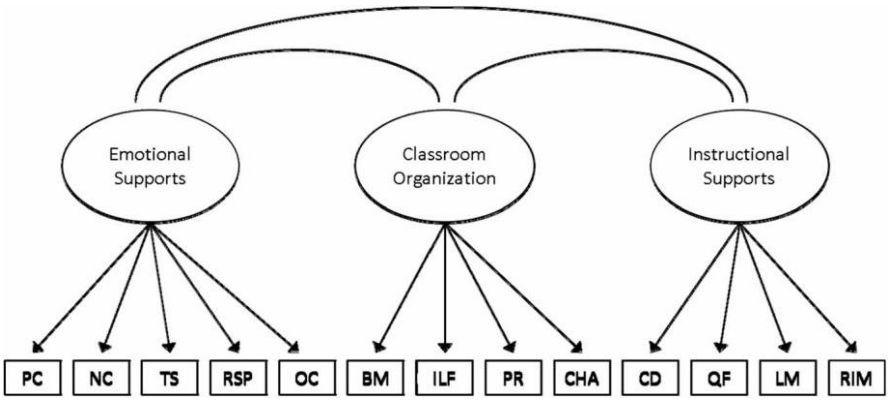

Figure 1. Teaching through Interactions factors (three-factor model). PC = Positive Climate; NC = Negative Climate; TS = Teacher Sensitivity; RSP = Regard for Student Perspectives; OC = Overcontrol; BM = Behavior Management; ILF = Instructional Learning Formats; PR = Productivity; CHA = Chaos; CD = Concept Development; QF = Quality of Feedback; LM = Language Modeling; RIM = Richness of Instructional Methods.

tween teachers and students or among students. Behavioral markers of relationships include physical proximity, shared activities, peer assistance, matched affect, and social conversation.

In the CLASS, these specific observable, behavioral descriptions are anchored at points along a seven-point rating scale to serve as a guide for observers' judgments regarding the quality of teacher-child interaction with regard to that dimension. These specific behavioral descriptions of dimensions and indicators shift across grades so that they are appropriate to that developmental level. The present study examines whether this three-domain latent structure of dimensions of teacher-child interactions applies in samples ranging from preschool through grade 5. The investigated samples are not limited to those that have specifically used CLASS, but instead encompass another, similarly scaled observational measure of teacher-child interactions (the Classroom Observation System [COS] used in the NICHD Study of Early Child Care and Youth Development; see NICHD Early Child Care Research Network, 2002, 2005).

It is important to note that this article does not focus on validating this framework against student outcomes. However, as discussed in detail below, there is extensive evidence supporting these links, most recently in the large MET study (Kane & Staiger, 2012). Additionally, although the CLASS measures many types of teacher-student interactions, it does not measure all of the possible elements of interactions that could be included within each broad domain described in the Teaching through Interaction framework. For example, Domitrovich et al. (2009) demonstrated that salient elements of the emotional environment, such as a teacher's use of emotion words and emotion coaching, were unique elements of classroom interactions not measured by the CLASS. Thus, although this study, based on the CLASS and COS measures, provides one way of testing the Teaching through Interactions framework, it could be tested using other measures and methods, including the use of other observational tools (Domitrovich et al., 2009) or student report data (Stuhlman, Downer, & DeCoster, 2012).

Table 1. Teaching through Interactions Framework: Description of Domains and Dimensions

Domain	Dimension	Description
Emotional Support	Positive Climate	Reflects the overall emotional tone of the classroom and the connection between teachers and students
	Negative Climate	Reflects overall level of expressed negativity in the classroom between teachers and students (e.g., anger, aggression, irritability)
	Teacher Sensitivity	Encompasses teachers' responsivity to students' needs and awareness of students' level of academic and emotional functioning
	Regard for Student Perspectives	The degree to which the teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations, and points of view, rather than being very teacher-driven
	Overcontrol	Assesses the extent to which the classroom is rigidly structured or regimented at the expense of children's interests and/or needs
Classroom Organization	Behavior Management	Encompasses teachers' ability to use effective methods to prevent and redirect misbehavior by presenting clear behavioral expectations and minimizing time spent on behavioral issues
	Productivity	Considers how well teachers manage instructional time and routines so that students have the maximum number of opportunities to learn
	Instructional Learning Formats	The degree to which teachers maximize students' engagement and ability to learn by providing interesting activities, instruction, centers, and materials
	Classroom Chaos	The degree to which teachers ineffectively manage children in the classroom so that disruption and chaos predominate
Instructional Support	Concept Development	The degree to which instructional discussions and activities promote students' higher-order thinking skills versus focus on rote and fact-based learning
	Quality of Feedback	Considers teachers' provision of feedback focused on expanding learning and understanding (formative evaluation), not correctness or the end product (summative evaluation)
	Language Modeling	The quality and amount of teachers' use of language-stimulation and language-facilitation techniques during individual, small-group, and large-group interactions with children
	Richness of Instructional Methods	The extent to which teachers use a variety of strategies to promote children's thinking and understanding of material at a deeper and more complex level

Theoretical and Empirical Background of the Teaching through Interactions Framework

Next we briefly describe the literature supporting the domains of teacher-child interactions that are the focus of this article. More detailed descriptions of the research supporting the specific dimensions of teaching practice are available elsewhere (e.g., Hamre & Pianta, 2007).

Emotional Support

Teacher efforts to support students' social and emotional functioning in the classroom, through positive facilitation of teacher-student and student-student interactions, are key elements of effective classroom practice. Two broad areas of developmental theory guide much of the work on emotional support in classrooms—attachment theory (Bowlby, 1969) and self-determination theory (Connell & Wellborn, 1991). Attachment theorists posit that when adults provide emotional support and a predictable, consistent, and safe environment, children become more self-reliant and are able to take risks as they explore the world because they know that an adult will be there to help them if they need it (Bowlby, 1969). Self-determination (or self-systems) theory (Connell & Wellborn, 1991) suggests that children are most motivated to learn when adults support their need to feel competent, positively related to others, and autonomous.

Classroom Organization and Management

The ways in which classrooms help students organize their behavior and attention toward the pursuit of academic goals is another salient domain of classroom interactions. Recent work demonstrates the importance of children's self-regulatory and executive functioning skills to learning and academic achievement (Blair, 2002; Ponitz, Rimm-Kaufman, Brock, & Nathanson, 2009). Classrooms that have clearer and more consistent routines for behavior and time use support children in developing these important skills (e.g., Brophy & Evertson, 1976; Emmer & Strough, 2001; Evertson, Emmer, Sanford, & Clements, 1983; Ponitz, Rimm-Kaufman, Grimm, & Curby, 2009). The strongest evidence for the importance of classroom organization and management comes from intervention studies. Children in classrooms in which teachers participate in interventions designed to enhance these aspects of their teaching demonstrate improvements in teacher reported and observed self-regulatory skills (e.g. Raver et al., 2009).

Instructional Support

The theoretical foundation for the conceptualization of instructional supports in the Teaching through Interactions framework comes primarily from research on children's cognitive and language development (e.g., Catts, Fey, Zhang, & Tomblin, 2001; Taylor, Pearson, Peterson, & Rodriguez, 2003). This literature highlights the distinction between simply learning facts and gaining "usable knowledge" that is built upon learning how facts are interconnected, organized, and conditioned upon one another (Mayer, 2002). A student's cognitive and language development is contingent on the opportunities adults provide to express existing skills and scaffold more complex ones (Skibbe, Behnke, & Justice, 2004). Furthermore, learning is strongest when teachers explicitly tie new information to students' background knowledge and real-world examples (Bransford, Brown, & Cocking, 2000). In addition to these elements of instructional interactions intended to enhance knowledge of concepts and language, effective teaching includes feedback that is immediate, contingent, corrective and/or specific, and tied to natural settings (e.g., Kulik & Kulik, 1988). Such feedback serves to control frustration; increase interest, motiva-

tion, and effort; and promote learning and higher-order thinking (Butler, 1987; Good & Brophy, 2008; Rogoff, 1990).

The Teaching through Interactions framework further differentiates between *general* and *content-specific* instructional supports. General instructional supports are those that are relevant and observable across content areas. Content-specific instructional supports, in contrast, describe strategies for teaching students particular skills and knowledge such as reading, math, or science. This study focuses on general instructional supports because these are the interactions that can be assessed most similarly across large numbers of classrooms. Numerous studies link these types of instructional interactions to students' academic learning (Catts et al., 2001; Hamre & Pianta, 2005; Howes et al., 2008; Taylor et al., 2003). However, we anticipate that many of the elements measured in observational tools such as the Mathematical Quality of Instruction (MQI; Hill et al., 2008) or the Protocol for Language Arts Teaching Observation (PLATO; Grossman et al., 2010), which focus on teachers' interactions with students around content, may be aligned with the larger Teaching through Interactions framework.

Teacher-Student Interactions Account for Student Learning and Developmental Gains

There is strong and consistent evidence that measures of classroom process informed by the Teaching through Interactions framework do indeed predict student performance. For example, observed Emotional Support predicts student performance in standardized tests of early literacy in preschool and first grade (NICHD Early Child Care Research Network, 2003), lower levels of internalizing behaviors reported by mothers in kindergarten and first grade (NICHD Early Child Care Research Network, 2003), and students' behavioral engagement with classrooms across several elementary grades (NICHD Early Child Care Research Network, 2002), and appears to help protect children at risk of school failure due to behavioral problems (Buyse, Verschueren, Doumen, Van Damme, & Maes, 2008; Gazelle, 2006). One recent study demonstrates that Emotional Support, as measured by CLASS, is associated with gains in standardized achievement and that these associations are mediated by student engagement (Reyes, Brackett, Rivers, White, & Salovey, 2012). Students also learn more in classrooms in which teachers do a better job managing students' behavior, time, and attention (Ponitz, Rimm-Kaufman, Brock, et al., 2009), partly due to children's development of better behavioral and cognitive self-control (Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009).

Finally, teachers' instructionally supportive interactions predict students' academic functioning (Hamre & Pianta, 2005; Howes et al., 2008) and behavioral engagement in classroom activities (NICHD Early Child Care Research Network, 2003). The MET study demonstrated that the CLASS measures aspects of effective teaching that are stable across class sections and that CLASS observations made in one class section were associated with gains in student achievement in other class sections (Kane & Staiger, 2012). Although a few studies have failed to demonstrate these associations (Strong, Gargani, & Hacifazlioğlu, 2011), the preponderance of evidence suggests that the Teaching through Interactions framework, as measured by the CLASS

and similar measures, describes aspects of the classroom process that contribute to student development and learning across assorted settings and grade levels.

Current Study

In this study, we examined a sample of 4,341 preschool to sixth-grade classrooms that were a part of seven national and regional studies covering a broad array of student and classroom characteristics. The observational data from these studies allow us to examine how a specific conceptual model of classroom settings can be applied to a broad spectrum of early childhood and elementary classrooms in the United States. To test the applicability and generalizability of the Teaching through Interactions framework's three-domain organization of teacher-child interactions, we first examined observational instruments used in these large-scale investigations and sorted observed dimensions of classroom process into the domains described by the Teaching through Interactions framework.

We then used confirmatory factor analysis to examine the extent to which the three-domain latent structure posited in the Teaching through Interactions framework (illustrated in Fig. 1) fits the natural variation in observed classrooms' processes in comparison with several alternative models. This type of analysis is important because the ways in which classroom interactions are organized has important implications for the understanding of classroom effects on student outcomes as well as for intervention work targeting improvement in these practices. The first alternative model (illustrated in Fig. 2) posits two domains, social and instructional supports. This model can be derived from a focus on classrooms that is isomorphic with the two most frequently assessed areas of student outcomes (achievement and social skills) and is consistent with some of the organizational frameworks that have been suggested in narrative reviews of classroom processes (Brophy, 1999). The second alternative model (illustrated in Fig. 3) posits a single domain of effective teaching, which is the underlying assumption of many teacher evaluation systems that use observational data to create a single score on teacher effectiveness.

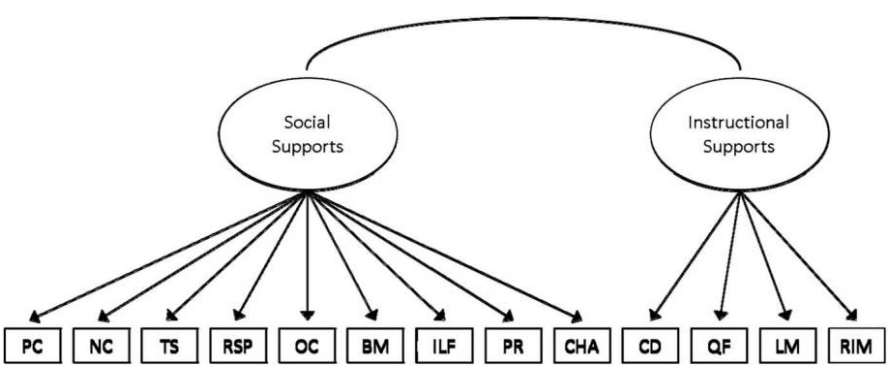

Figure 2. Social and instructional factors (two-factor model). PC = Positive Climate; NC = Negative Climate; TS = Teacher Sensitivity; RSP = Regard for Student Perspectives; OC = Overcontrol; BM = Behavior Management; ILF = Instructional Learning Formats; PR = Productivity; CHA = Chaos; CD = Concept Development; QF = Quality of Feedback; LM = Language Modeling; RIM = Richness of Instructional Methods.

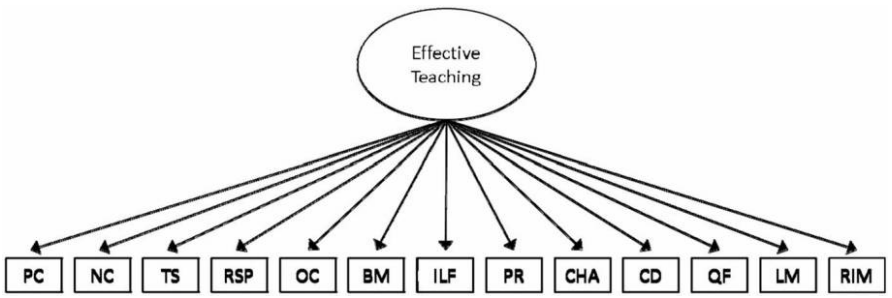

Figure 3. Effective teaching factor (one-factor model). PC = Positive Climate; NC = Negative Climate; TS = Teacher Sensitivity; RSP = Regard for Student Perspectives; OC = Overcontrol; BM = Behavior Management; ILF = Instructional Learning Formats; PR = Productivity; CHA = Chaos; CD = Concept Development; QF = Quality of Feedback; LM = Language Modeling; RIM = Richness of Instructional Methods.

Finally, we examined the extent to which the three-domain model held true across the 10 observational studies representing classrooms from preschool to sixth grade. Consistent with the theoretical and empirical data cited above, we expected that the three-domain model, in which dimensions were organized under Emotional, Organizational, and Instructional interactions, would provide a better fit to the data than either of the two alternative models.

Method

Studies and Sample

The present study utilizes data from seven large-scale observational research projects conducted from 1998 to 2009 in 4,341 preschool to sixth-grade classrooms across the United States. Basic information for each study's sample and the classroom observations is included below and in Table 2. A description of each of the samples is provided below. Readers are referred to individual study citations for more complete information on the data-collection procedures and sample.

National Center for Early Development and Learning Multi-State Study of Prekindergarten (NCEDL MS). The NCEDL MS (Pianta et al., 2005) was conducted in six states with state-funded preschool programs. In each state (or large metropolitan area within the state), a stratified random sample of 40 preschool centers or schools was selected from a list provided by the state's department of education. Of the initial sites that were eligible, 78% agreed to participate, resulting in 240 prekindergarten classrooms enrolled across the 2001–2002 academic year. These children were then followed into 737 kindergarten classrooms, which also were included in these analyses. In NCEDL MS pre-K classrooms, 24 observers made two classroom visits in the fall and two in the spring. Observers coded for an entire day in half-day programs and until nap time in full-day programs, resulting in an average of 6.03 30-minute cycles per visit (SD = 1.49). In NCEDL MS kindergarten classrooms, 24 observers made three classroom visits spaced throughout the year. These observations occurred for an entire school day, except for recess, lunch, and nap, resulting in an average of 7.70 30-minute cycles per visit (SD = 2.34).

Table 2. Descriptive Data for Classroom, Teacher, and Child Demographics and Observation Procedures

	Study									
	NCEDL Multi-State		NCEDL SWEEP	MTP	NICHD-SECCYD	NICHD-SECCYD	NICHD-SECCYD	NYC Study	Links to Learning	RULER
	Pre-K	K	Pre-K	Pre-K	Grade 1	Grade 3	Grade 5	Grades 3–4	K–Grade 4	Grades 5–6
Classrooms:										
Number observed	240	737	454	152	834	827	791	152	61	93
Class size—*M* (*SD*)	18.12 (3.53)	20.39 (3.74)	17.31 (4.91)	14.44 (2.45)	19.25 (4.68)	18.99 (4.35)	20.77 (5.30)	18.53 (5.69)	20.27 (6.92)	25.30 (5.48)
Teachers:										
Education:										
High school or less	4	–	1	–	–	–	–	–	–	–
AA/AS or some college	37	2	22	–	–	–	–	.7	–	–
BA degree	33	52	55	66	57	55	49	30	61	53
Advanced degree	26	43	21	34	38	41	46	69	34	45
Years of experience—*M* (*SD*)	13.17 (9.35)	14.18 (9.81)	13.37 (9.25)	15.04 (9.45)	14.34 (9.43)	13.87 (10.48)	14.74 (10.86)	6.49 (6.15)	12.90 (12.34)	14.50 (10.03)
Race/ethnicity										
Non-Hispanic White	60	73	67	70	90	88	88	52	39	81
African American	17	7	7	24	3	4	4	29	41	11
Hispanic	12	10	14	<1	1	2	1	15	3	9
Other	11	10	12	5	6	4	7	4	5	–
Children:										
Gender (% female)	48	48	48	52	51	51	52	51	49	50
Maternal education—*M* (*SD*)	12.41 (1.18)	12.53 (2.16)	13.01 (1.44)	12.71 (1.04)	14.48 (2.44)	14.52 (2.42)	14.57 (2.33)	12.58[a] (2.37)	12.26 (1.80)	–
Household income—*M* (*SD*)	28,733[b] (15,457)	30,202[b] (22,686)	34,805[b] (20,013)	1.36[c] (.56)	4.01[c] (3.08)	4.40[c] (3.60)	4.68[c] (4.06)	1.71/1.91 (1.71/2.02)	.70[d] (1.03)	–

Race/ethnicity:										
White	41	47	41	26	84	83	83	4	0	23
African American	22	24	15	49	11	11	11	42	98	33
Hispanic	26	23	28	12	4	4	4	46	0	29
Other	11	6	16	13	1	2	2	8	2	13
Observation procedure:										
Instrument	CLASS	CLASS	CLASS	CLASS	COS-1	COS-3	COS-5	CLASS	CLASS	CLASS
Year of observation	2001–2002	2002–2003	2003–2004	2004–2005	1998	2000	2002	2005–2006	2006	2008
Number of days in classroom—*M* (*SD*)	4.47 (1.03)	2.86 (.59)	1.33 (.53)	8.10 (2.11)	1	1	1	1–3	1	3
Number of cycles per day—*M* (*SD*)	6.03 (1.49)	7.70 (2.34)	6.84 (1.84)	1	2.93 (.26)	7.96 (.24)	7.92 (.33)	3.9 (.29)	4	2
Cycle length in minutes	30	30	30	24	15	15	15	20	20	15
Method of observation	live	live	live	videotape	live	live	live	live	live	videotape
Number of observers	24	24	20	13	36	36	37	21	12	10
Reliability:										
Weighted kappa (range)	.70 (.42–.83)	.64 (.50–.82)	.58 (.52–.66)	–	–	–	–		–	
Mean % within 1 (range)	.91 (.80–1.00)	.86 (.80–.98)	.85 (.80–.90)	.85 (.75–1.00)	.71 (.55–.87)	.71 (.45–.89)	.87 (.72–.95)	.86 (.76–.95)	.91 (.86–1.00)	.89 (.85–.93)

Note.—Some percentages do not add up to 100 due to missing data. NCEDL = National Center for Early Development and Learning; SWEEP = State-Wide Early Education Programs Study; MTP = MyTeachingPartner; NICHD-SECCYD = National Institute of Child Health and Human Development Study of Early Child Care and Youth Development.

[a] Highest education level of any adult in the household.

[b] Total household income.

[c] Federal income-to-needs ratio.

[d] 0 = 0–$10,000; 1= $11,000–$20,000; 2 = $21,000–$30,000; 3 = $31,000–$40,000; 4 = $41,000–$50,000; 5 = over $50,000.

NCEDL State-Wide Early Education Programs Study (NCEDL SWEEP). The NCEDL SWEEP study was conducted in five states that were selected to complement those included in the NCEDL MS sample, mainly to diversify funding and delivery models. This study set out to randomly select 100 preschool centers from each state (or regions within states for those states with large metropolitan areas), resulting in a total of 454 classrooms enrolled in the study across the entire 2003–2004 academic year. During one classroom visit in the spring, 20 observers coded for an entire day in half-day programs and until nap time in full-day programs; this resulted in an average of 6.84 30-minute cycles per visit ($SD = 1.84$).

MyTeachingPartner (MTP). MyTeachingPartner (Pianta, Mashburn, Downer, Hamre, & Justice, 2008) was an NICHD-funded professional development project targeting the quality of pre-K teachers' classroom interactions with students. From one state-funded pre-K program, 240 teachers were recruited to participate across 41 school districts. Videotapes of classroom interactions were available for 152 of these teachers during the 2004–2005 academic year and were included in the current study. Teachers were provided with a video camera and tripod and asked to videotape themselves for 30 minutes once every 2 weeks. Teachers were asked to tape their implementation of either a language/literacy or socioemotional curriculum (58% language/literacy) and were requested to tape 5 minutes prior to the lesson as well as anything after the lesson up to 30 minutes. These videotapes were then coded using CLASS by a team of 13 observers. A small sample of tapes was double coded for reliability purposes.

National Institute of Child Health and Human Development Study of Early Child Care and Youth Development (NICHD SECCYD). The NICHD SECCYD (NICHD Early Child Care Research Network, 2002, 2005; Pianta, Belsky, Houts, Morrison, & NICHD Early Child Care Research Network, 2007) was a longitudinal study of key developmental contexts for children from birth through eighth grade. Families were recruited through hospital visits to mothers shortly after the birth of a child in 1991 at 10 locations in the United States. 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 present study included data from participating children in first-grade ($N = 834$), third-grade ($N = 827$), and fifth-grade ($N = 791$) classrooms. Classroom observations took place in winter to early spring of first, third, and fifth grades. Within NICHD SECCYD first-grade classrooms, observations occurred during the morning, began with the official start of the school day, and lasted approximately 3 hours. Within third- and fifth-grade classrooms, observations occurred for an entire school day, except for recess and lunch (approximately 6 hours). Observations in all three grades took place on a single day and consisted of eight 15-minute cycles.

New York City Study of Social and Literacy Development. The New York City Study of Social and Literacy Development (Brown, Jones, LaRusso & Aber, 2010; Jones, Brown, & Aber, 2011) was a 3-year longitudinal, experimental evaluation of a universal, whole-school intervention (the 4Rs Program: Reading, Writing, Respect, and Resolution). The evaluation of 4Rs was conducted in 18 NYC public elementary schools. Live classroom observations were completed in 152 third-, fourth-, and fifth-grade classrooms by a total of 15 observers in all intervention and control schools. Blind to school intervention status, observers scheduled observations in 2-hour blocks during regular "instructional" time. Observers conducted four 20-minute

observational segments, each followed by a 10-minute coding segment. The majority of classrooms were observed only once, though in a few circumstances observers needed to return for a second or third day due to irregularities such as a change in schedule.

Links to Learning study. Links to Learning is an intervention study in high-poverty urban schools examining the feasibility and effectiveness of an intervention to promote learning among children with disruptive behavior problems (Cappella, Frazier, Atkins, Schoenwald, & Glisson, 2008). Schools were selected for the first cohort of the intervention trial based on a set of demographic criteria to ensure their comparability and facilitate collaboration with social service agencies. Baseline data from 61 K through fourth-grade regular education classrooms across the six Cohort 1 schools were included in this sample. Classroom observations were conducted live in the fall of 2006 by 12 observers. Observations began at the start of the school day and lasted for 2 hours of morning instruction with the lead teacher. Teacher-student interactions were observed for 20 minutes each, four times during the 2-hour period. Following each 20-minute interval, the observer spent 10 minutes recording scores for each of the CLASS dimensions. Although most observations occurred during one morning period, a small number of classrooms required a second observation day to complete the four intervals.

RULER Intervention study. The RULER study (Brackett, Rivers, Reyes, & Salovey, 2010) was a large-scale randomized controlled trial that evaluated the effectiveness of a language-based social and emotional learning program in fifth- and sixth-grade English language arts (ELA) classes from a Catholic school district in Brooklyn and Queens, New York. Sixty-four schools volunteered to participate in the study, including 155 ELA classes. At the baseline (prior to randomization, spring 2008), more than half of the teachers (56%) submitted videotapes of their classroom practice. This subsample comprises 93 classrooms (68 teachers) with 1,903 students across 46 schools. Teachers were provided with a video camera and asked to videotape their entire ELA class on 3 separate days. Each of these videotapes was approximately 30 minutes and was separated into two cycles; one cycle included the first 15 minutes of footage of the lesson, and the other cycle included the next 15 minutes of footage of the lesson (average segment length = 14.8 minutes). Though rare, shorter cycles were considered viable and included in the study if they were 8 minutes or longer. Each cycle was coded at least once by one of 10 observers, and 40% of cycles were coded by four observers. On average, 5.35 cycles per classroom and 1.86 cycles per lesson within a classroom on a given day were coded.

Observational Measures

One of two observational systems was used to conduct live or videotaped observations in each of the projects: the Classroom Assessment Scoring System (CLASS) or its precursor, the Classroom Observation System (COS). More recent studies used the CLASS, whereas older studies used the COS. The individual scales measured by each of these systems are described in Table 2. The CLASS was developed using the COS as a guide, and the two observational systems share a conceptual framework and scoring protocol.

Classroom Assessment Scoring System (CLASS; Pianta, La Paro, et al., 2008). The CLASS was developed to assess classroom quality from pre-K to high school

(see www.classobservation.com); however, the versions of the CLASS used in this report focused on pre-K to fifth grade. An early version of the CLASS used in the NCEDL studies contained nine dimensions, whereas the current version used in MTP, RULER, the New York City Study of Social and Literacy Development, and Links to Learning studies contained 10. In each study, global ratings during each cycle (ranging from 15 to 30 minutes) were made on a 7-point scale, ranging from "low" to "high," for the following dimensions of teacher-student interactions: positive climate, negative climate, teacher sensitivity, regard for student perspectives, behavior management, instructional learning formats, productivity, concept development, quality of feedback, and language modeling. One dimension from the CLASS used in the NCEDL studies, overcontrol, was significantly revised due to problems with limited variability and skewness into a new dimension, Regard for Student Perspectives. Observers watched classroom interactions for a prescribed period of time (which varied slightly by study, but ranged from 15 to 30 minutes) while taking detailed field notes about specific teacher and student behaviors and interaction patterns. Observers then had 10 minutes to use the behaviorally anchored set of rating scales in the CLASS manual to determine a final code for each dimension. CLASS scores used in current analyses were aggregated across cycles, observers, and observation visits to form variables at the classroom level.

COS-1, -3, -5. The first-, third-, and fifth-grade versions of the COS were developed specifically for use in the NICHD SECCYD to track child behaviors and classroom conditions. The COS uses a multilevel observation format that incorporates both discrete codes and global ratings, although only the global ratings of classroom quality were used in this study. Ratings during each cycle were made on a 7-point scale ranging from "uncharacteristic" to "extremely characteristic" for the following dimensions of teacher-student interactions: literacy instruction, evaluative feedback, instructional conversation, encouragement of child responsibility, positive emotional climate, negative emotional climate, classroom management, teacher sensitivity, richness of instructional methods, and classroom chaos (see Table 3). Observers watched classroom interactions for 5 minutes prior to a 10-minute time-sampling period, as well as 10 minutes immediately after the time-sampled period. During this dedicated observation time, observers took detailed field notes about specific teacher and student behaviors and interaction patterns. Observers then had 10 minutes to use the behaviorally anchored set of rating scales in the COS manual to determine a final code for each dimension. COS scores used in current analyses were aggregated across cycles to form variables at the classroom level.

Training and reliability. Across studies, all observers attended a centralized workshop to attain reliability on the CLASS or COS. Prior to attending the workshop, observers were asked to read a manual with extensive descriptions of dimensions and anchor points and to practice coding several videotape clips. The workshops consisted of guided practice with coding videotaped classroom footage. After the training workshops, observers had to pass a reliability test involving the coding of either five or six cycles of 20–44 minutes. Observers' global ratings had to match the master global ratings (within 1 scale point) on at least 80% of the cycles for them to pass the test. During data collection, all studies completed further reliability checks through independent live double-coding or

Table 3. Means, Standard Deviations, and Correlation Matrix among Teacher-Student Interaction Variables ($n = 458$ to $n = 4{,}341$)

	PC	NC	TS	RSP	OC	BM	ILF	PR	CH	CD	QF	LM	RIM
Mean	5.10	1.40	4.75	4.26	2.12	5.26	4.14	4.82	1.57	2.58	2.54	2.81	2.80
SD	.803	.61	.90	.88	1.09	.91	1.02	.94	.67	1.02	1.13	.93	1.12
PC	1.00												
NC	−.56	1.00											
TS	.76	−.56	1.00										
RSP	.53	−.36	.68	1.00									
OC	−.48	.41	−.31	−.46	1.00								
BM	.54	−.51	.61	.41	−.31	1.00							
ILF	.44	−.29	.49	.66	−.21	.51	1.00						
PR	.47	−.44	.53	.33	−.09	.70	.61	1.00					
CH	−.32	.44	−.28		−.02			−.54	1.00				
CD	.21	−.21	.30	.42	−.02	.32	.45	.45		1.00			
QF	.32	−.22	.38	.38	.05	.30	.43	.45	−.23	.78	1.00		
LM	.41	−.28	.65	.65	−.32	.31	.43	.21		.75	.54	1.00	
RIM	.33	−.01	.25	.25	−.37			.29	−.19		.64		1.00

Note.—All correlations above .05 are significant at $p < .001$. Missing correlations could not be estimated as these variables were not present in the same data set. PC = Positive Climate; NC = Negative Climate; TS = Teacher Sensitivity; RSP = Regard for Student Perspectives; OC = Overcontrol; BM = Behavior Management; ILF = Instructional Learning Formats; PR = Productivity; CH = Chaos; CD = Concept Development; QF = Quality of Feedback; LM = Language Modeling; RIM = Richness of Instructional Methods.

independent dual review of videotapes. Assessments of reliability within each study, as available, are reported in Table 2.

Data Analysis Approach

Data organization and analysis involved three steps: content analysis, confirmatory factor analysis, and multiple group confirmatory factor analysis. First, based on literature reviewed in the introduction, six members of our research team independently sorted CLASS and COS dimensions twice. First they were sorted into the hypothesized three domains from the Teaching through Interactions framework. Then they were sorted into two domains: Social Supports and Instructional Supports. In both cases, there was 100% agreement on the domain in which each dimension measured by CLASS or COS best fit. The final model that was tested placed all dimensions into a single factor. Figures 1, 2, and 3 provide a summary of each of these models.

We then conducted a series of confirmatory factor analyses (CFAs) to determine whether the Teaching through Interactions model (with three factors), the Social and Instructional Supports model (with two factors), or the Effective Teaching model (with one factor) best fit the observed data. These models were examined both in the full data set including all of the samples, as well as in the data set for each individual sample. Multiple-fit indices were then compared across the three models because the chi-square overall goodness-of-fit test is unfavorably affected by large sample size, model misspecification, or violation of distribution assumptions (Bollen, 1990). Given that our samples all had missing data, we used multiple imputation, a method that has been identified as one of the optimal ways to handle missingness in educational research (Peugh & Enders, 2004), when estimating and testing our models.

Analyses were conducted and aggregated across 100 imputed data sets to ensure that our results represented valid, consistent patterns in the data. Bayesian analysis was used to generate the imputed datasets.

Results

Descriptive Characteristics

Table 3 shows descriptive statistics including means, standard deviations, and correlation coefficients among observed classroom interaction variables. The correlations presented in Table 3 demonstrate some areas of high convergence (e.g., strong association between Positive Climate and Teacher Sensitivity) and divergence (e.g., low correlations between Overcontrol and Concept Development).

CFA Models

Figures 1–3 depict path diagrams for three alternative structural equation models that would explain the factor structure of our interaction measures. Our main interest was in the Teaching through Interactions model (illustrated in Fig. 1), where the dimensions were assigned to emotional supports, classroom organization, and instructional supports factors. We also tested two additional models providing alternative organizations of the interaction measures. In the Social/ Emotional Supports model (illustrated in Fig. 2), we collapsed Emotional Support and Classroom Organization factors to form a single latent construct, isolating Instructional Support. In the Effective Teaching model (illustrated in Fig. 3), all of the measures loaded on a single factor. Mplus 6.11 was used to fit all three models using maximum-likelihood estimation.

Table 4 shows the fit indices for the three CFA models when they were fit to the complete data set. According to Hooper, Coughlan, and Mullen (2008), RMSEA values less than .05 indicate good fit and values less than .10 indicate adequate fit, CFI values greater than .95 indicate good fit and values greater than .90 indicate adequate fit, and TLI values greater than .95 indicate good fit. Based on these cutoffs, the Teaching through Interactions and the Social/Instructional Supports models showed evidence of good fit, while the Effective Teaching model showed evidence of adequate fit. The CFI and TLI statistics suggest that all three models have less than adequate fit. Across all of the

Table 4. Fit Statistics for the Different CFAs in the Combined Sample

	Teaching through Interactions	Social and Instructional Supports	Effective Teaching
Factors	3	2	1
χ^2	728	849	993
df	62	64	65
CFI	.844	.816	.782
TLI	.803	.776	.739
RMSEA	.047	.050	.054
95% CI	(.044, .050)	(.047, .053)	(.051, .057)
SRMSR	.045	.049	.053

Note.—CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMSR = standardized root mean square residual.

Table 5. Factor Loadings for the Teaching through Interactions Model

	Emotional Support		Classroom Organization		Instructional Support	
Dimensions	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Positive Climate	.834	.011				
Negative Climate	−.631	.013				
Teacher Sensitivity	.560	.017				
Respect for Student Perspective	.221	.020				
Overcontrol	−.441	.015				
Behavior Management			.512	.019		
Instructional Learning Format			.329	.022		
Productivity			.532	.022		
Chaos			−.139	.025		
Concept Development					.346	.026
Quality of Feedback					.674	.034
Language Modeling					.237	.024
Richness of Instruction					.159	.025

fit statistics, the Teaching through Interactions model shows better fit than the Social/Instructional Supports and Effective Teaching models.

The standardized coefficients for the Teaching through Interactions model are presented in Table 5. All of the coefficients were significant ($p < .001$), indicating that each of the measures was significantly related to its factor. Correlations among the factor scores indicated that all of the factors were significantly related to each other (all p's $< .001$). The associations of Classroom Organization with Emotional Support ($r = .75$) and Instructional Support ($r = .59$) were strong effects. The association between Emotional Support and Instructional Support was lower ($r = .43$) but still between a medium and large effect.

Reliability scores were computed for each of the three factors within the Teaching through Interactions model. The formula for the reliability of the composite score is defined for the confirmatory factor model as

$$\rho_{xx} = \frac{\sum_{i=1}^{m} \lambda_i^2 + \sum_{i}^{m} \sum_{j}^{m} \lambda_i \lambda_j}{\sum_{i=1}^{m} \lambda_i^2 + \sum_{i}^{m} \sum_{j}^{m} \lambda_i \lambda_j + \sum_{i}^{m} \theta_{ii}} \qquad i \neq j,$$

where m is the number of indicators and λ_I and θ_I are a factor loading and residual variance for the ith variable, respectively. The reliability of Emotional Support was $\rho_{xx} = .68$, the reliability of Classroom Organization was $\rho_{xx} = .78$, and the reliability of Instructional Support was $\rho_{xx} = .93$. These internal consistency measures are close to or greater than an acceptable cutoff value of $\rho_{xx} = .70$.

Summary of model fit statistics for tests of the three different CFA models within each data set (after imputation) is shown in Table 6. We were unfortunately unable to test the fit of the Teaching through Interactions and Social/Instructional Supports models in two of our samples (NICHD SECCYD first grade and NICHD SECCYD third grade) because they did not have multiple measures of Instructional Support. However, across all of the other samples, the Teaching through Interactions model had consistently better fit statistics across all of the samples than the Social/Instructional Supports and Effective Teaching models.

Table 6. Summary of Model Fit Indices as a Function of Study Samples and Alternative Confirmatory Factor Models

Sample	Classroom *N*	Teaching through Interactions Model (*df* = 62)			Social/Instructional Supports (*df* = 64)			Effective Teaching (*df* = 65)		
		χ^2	CFI	RMSEA	χ^2	CFI	RMSEA	χ^2	CFI	RMSEA
NCEDL Multi-State Pre-K	240	110	.95	.057	165	.90	.081	248	.82	.108
NCEDL SWEEP Pre-K	454	153	.94	.057	230	.88	.075	379	.78	.103
MTP Pre-K	152	80	.97	.044	100	.95	.061	147	.88	.091
NCEDL KG	737	273	.92	.068	364	.89	.08	721	.76	.117
NICHD- SECCYD Grade 1	834	–	–	–	–	–	–	114	.96	.024
NICHD- SECCYD Grade 3	827	–	–	–	–	–	–	259	.83	.060
NICHD- SECCYD Grade 5	791	284	.84	.067	755	.55	.117	520	.71	.094
NYC Study	152	103	.96	.066	133	.93	.084	192	.88	.114
Links to Learning	61	65	.99	.027	91	.91	.083	104	.86	.100
RULER	93	139	.80	.116	160	.76	.127	193	.67	.146

Note.—Fit indices could not be computed within the NICHD–SECCYD first- and third-grade samples for the Teaching through Interactions or the Social/Instructional Supports models. NCEDL = National Center for Early Development and Learning; SWEEP = State-Wide Early Education Programs Study; MTP = MyTeachingPartner; NICHD-SECCYD = National Institute of Child Health and Human Development Study of Early Child Care and Youth Development.

Discussion

Education research, policy, and practice are faced with the daunting task of unpacking the "black box" of teaching—what is it that teachers do on a daily basis that contributes to students' development of academic and social competencies? This is a question of central importance in the current policy context where "teacher effectiveness" is a key part of federal education initiatives, such as Race to the Top, and a part of many teacher evaluation and compensation programs. Although there is an abundance of research identifying specific types of instructional strategies that are effective in particular contexts, there is a need for clearer articulations and validation of higher-level theories of effective teaching (Douglas, 2009).

As exemplified in the recent MET study, there are numerous approaches to conceptually organizing and measuring the elements of teaching practice that drive student learning. The MET study included five different observational measures, two general measures and three content-specific measures, which each posit different sets of classroom practices. For example, Charlotte Danielson's Framework for Teaching (Danielson, 2007) suggests four domains of effective teaching—Planning and Preparation, Professional Responsibilities, Classroom Environment, and Instruction—each of which is composed of multiple components. However, there are not any peer-reviewed, published papers testing the extent to which this theorized factor structure conforms to the reality of classrooms. As discussed in greater detail below, the ways in which we come to understand how components of teaching are organized into broader domains have important research and practice implications. As such, it is important to validate any of these theoretical models using data from diverse

samples. This study provides initial evidence of the validity of one such approach, the Teaching through Interactions framework.

The Teaching through Interactions framework posits that much of the effect of teachers and classrooms on student learning is located in the interactions that take place between teachers and students. This model suggests that across grade levels, these interactions are organized into three domains—Emotional Support, Classroom Organization, and Instructional Support. A growing body of research documents the connections between these types of interactions and student outcomes in preschool through sixth grade (Buyse et al., 2008; Gazelle, 2006; Hamre & Pianta, 2005; Kane & Staiger, 2012; NICHD ECCRN, 2003; Pianta et al., 2007; Reyes et al., 2012; see Strong et al., 2011, for exception). The present study extends this validity work by providing evidence that this organization of classroom interactions into three broad domains of effective teaching "fits" the reality of preschool and elementary classrooms across the country. When tested against models suggesting either an omnibus teacher quality factor or a simpler social and instructional supports model, the three-factor model fit data from 4,341 early childhood and elementary classrooms the best. This three-factor model fit observational data collected from a range of studies, using different measures and observational methodologies (e.g., live vs. video), across a broad range of settings, including urban and rural classrooms, and across preschool to sixth-grade classrooms.

Below we discuss two broad areas of research on effective teaching that may be enhanced by the use of the more theoretically driven and empirically supported frameworks for teaching, such as the Teaching through Interactions framework: the development of effective teaching and understanding associations between teaching and students' learning.

Understanding the Development of Effective Teaching

One way in which the articulation of three major domains of effective teaching may be helpful is in guiding research on ways in which teachers develop specific skills over time. For example, Malmberg, Hagger, Burn, Mutton, and Colls (2011) followed teachers from their last year of teacher education, to a student teaching placement, and then into their first 2 years of teaching. The results showed that, on average, teachers increased their classroom organization and management skills over these early years of teaching, with the strongest improvements seen among teachers who started off relatively lower in this domain. A different pattern was found for emotional support, where initial increases were followed by declines, leading to an inverted U-shape curve over time. Malmberg and colleagues suggested that this finding may reflect stress and demands on beginning teachers as they move from teacher education to their own jobs. This research offers a more nuanced view of the development of teaching skills than do studies that simply examine changes in teacher performance over time based on student test scores or teacher report, or those that examine one dimension of teaching at a time.

Recent work also suggests that these domains of teaching may be improved through intentionally designed and intensive professional development experiences (e.g., Brown et al., 2010; Domitrovich et al., 2009; Hamre et al., 2012; Pianta, Mashburn, et al., 2008; Raver et al., 2008). Many of these studies show differential impacts of professional development on particular domains of teaching. For example, Brown

and colleagues (2010) demonstrated differences between classrooms implementing the 4Rs social-emotional and literacy curriculum and control classrooms implementing Emotional and Instructional Supports but not Classroom Organization. Similar results were obtained in a study testing the efficacy of a course designed to enhance teacher-student interactions among early childhood teachers (Hamre et al., 2012). An intervention focused on using mental health consultants in classrooms demonstrated impacts on only Emotional Support (Cappella et al., 2012). In contrast, other studies have provided evidence that Classroom Organization is sensitive to intervention (Domitrovich et al., 2009; Raver et al., 2008). These types of studies can help us better understand the specific professional development experiences that are most likely to impact particular domains of teaching.

A final literature of relevance here examines predictors of effective teaching. Rimm-Kaufman and Hamre (2010) proposed a Comprehensive Model of Teacher Quality that embeds effective teaching in a dynamic model that considers the ways teachers' personal attributes and professional experiences influence the development of their teaching skills, as well as the ways in which these relationships are influenced by contextual factors such as professional development, school climate, and district policies. This model recognizes the complex systems that influence the development of teaching skills and can be made more complex—and more precise—by a consideration of the ways these systems operate differently in influencing the three broad domains of teaching. For example, within early childhood literature there are few indications that teacher degree status influences effectiveness (Early et al., 2007); however, to the extent this evidence exists, it is primarily in the instructional domain (LaParo et al., 2009; Pianta et al., 2005). The lack of an association between teacher education and classroom organization or emotional support suggests that current teacher education programs do very little to teach teachers how to develop positive relationships with students or develop effective management strategies. In contrast, teachers' psychological characteristics appear to be more closely tied to emotional rather than instructional supports (Li-Grining et al., 2010), whereas teacher experience is most closely linked to skills in management and organization (La Paro et al., 2009). Again, the point here is that by conceptualizing and measuring multiple domains of teaching practice, we are better able to understand the factors that influence teachers' development of the complex set of skills that are required to be an effective teacher.

Understanding Associations between Teaching and Student Development

Another area of research in which the Teaching through Interactions framework may enhance our knowledge is the effects of teaching on student learning and development. There is an assumption underlying the theoretical basis for this framework of domain specificity: emotional supports promote social development, classroom management and organization promote positive behavior and attention, and instructional supports enhance learning. Although some research supports this domain specificity (e.g., Howes et al., 2008; Mashburn et al., 2008; Rimm-Kaufman et al., 2009), the developmental theory on which this framework was based suggests that these effects are not so simple (Downer, Sabol, & Hamre, 2010). There are some examples from the literature supporting cross-domain linkages. For example, there are several studies indicating that emotional support is associated with student learn-

ing (Hamre & Pianta, 2005; Rudasill, Gallagher, & White, 2010). This is important because in the current context of accountability for student achievement, there is a tendency for schools to focus only on instruction. This research suggests that attending to teachers' emotional supports may also be important, especially for children with social and behavioral problems (Hamre & Pianta, 2005). Research of this type is more convincing if it does not focus solely on emotional supports, but rather compares effects across these different domains of teaching.

Other Approaches to Understanding Teacher Effectiveness

Although the Teaching through Interactions framework offers one broad, validated approach to understanding teacher effectiveness, there are clearly many other components to effective teaching, both within and outside the broad domain of teacher-student interactions. An important part of the Teaching through Interactions framework, but one that was not measured by either the CLASS or COS, is the idea of content-specific instructional supports. For example, recent observational work has focused on the specific instructional strategies used in successful mathematics teaching (Hill et al., 2008) and English language arts instruction (Grossman, 2011). In the MET study (Kane & Staiger, 2012), these content-specific instruments were moderately to highly correlated with more general measures of teaching practice, such as CLASS and the Danielson Framework for Teaching, suggesting there are many elements of teaching that may be shared across content areas, but also that each content area may have unique elements of instruction. Within the MET study, the general measures of teaching practice predicted gains in student achievement at similar levels as did the more content-specific measures. It will be important for future work to examine how these types of content-specific instructional supports may fit with more general instructional strategies. Do they form a unique component of effective teaching, as hypothesized, or are they well aligned with other more general instructional supports offered by teachers?

It is also important to note that although observation offers an important window into teaching, it is not the only methodology to do so. As one example, teacher logs provide a cost-effective method for gathering data about teacher practices across the year (Rowan & Correnti, 2009). Although much of the initial research using teacher logs focused primarily on the amount of instruction provided across a variety of curricular activities, more recent work has also captured more qualitative components of teaching, such as the extent to which teachers use strategies that provide a higher cognitive demand (Rowan & Correnti, 2009). Student reports on teachers offer another important window on teacher effectiveness. MET study results suggest that student reports are often more highly correlated with teacher value-added scores than are observational data (Kane & Staiger, 2012).

There is also a great deal that teachers do that falls outside of their interactions with students. For example, they create assignments for students that can be assessed based on dimensions such as cognitive challenge, clarity of learning goals, clarity of grading criteria, and alignment of goals and task (Clare & Aschbacher, 2001). There is evidence that the quality of these assignments is related to student learning, above and beyond observational measures of teachers' practice (Matsumura, Garnier, Slater, & Boston, 2008). Assessments of teacher effectiveness should also attend to these components of teachers' jobs.

The Teaching through Interactions framework offers only one window into teacher effectiveness. However, we argue that interactions among teachers and students are among the most important aspects of teachers' jobs. There is also initial evidence that this model for understanding classroom interactions is consistent across other cultures (Cadima, Leal, & Burchinal, 2010; Malmberg & Hagger, 2009; Pakarinen et al., 2010) and into secondary settings (Allen, Pianta, Gregory, Mikami, & Lun, 2011; Malmberg & Hagger, 2009). Furthermore, observations can be conducted as a part of normative school practice (e.g., principle observations for evaluation purposes) and can work across contexts, unlike methods such as student report that would be difficult to complete with younger students.

Limitations and Future Directions for Research

There are a few other limitations to note beyond those already discussed. First and foremost, this study tested a conceptual model of teaching using two observational systems that were developed, at least in part, on this theory. Theory cannot, however, move forward based on a single measure. There is a need for future work to examine the extent to which data from other measures, based on different frameworks but also focusing on teacher-student interactions (e.g., Connor et al., 2009; Matsumura et al., 2008), may also fit this broad conceptual model for the way teacher-student interactions are organized. There are also many elements of interaction neither sufficiently covered by this theory nor measured by the CLASS or COS. Beyond the content-specific elements of teacher-student interactions discussed above, there is a need to attend to interactions around cultural sensitivity and social justice (Kaur, in press). With regard to the measures used in this study, the interrater reliabilities for the global rating scales used in these studies, while adequate, were not as strong as might be desirable, reflecting the presence of unexplained influences on the scores.

Additionally, the covariance among the three domains of teacher-student interactions are high, suggesting that there are significant elements of effective teaching that are shared across domains. These results are consistent with the MET study, which suggested a single overall factor for the CLASS (Kane & Staiger, 2012). A high level of covariance across domains can lead to analytic challenges when these domains are used together to predict outcomes (Downer et al., 2010). One potential analytic approach that could be used to address this challenge is the bifactor approach (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012). Bifactor modeling differs from traditional factor analysis in that it allows for general qualities between the factors to be estimated as a general factor as well as estimating specific factors over and beyond the general factor. Once these factors are estimated, they are uncorrelated and thus represent unique features of a multifaceted construct. Future work with CLASS may examine the extent to which this approach may offer both analytic and conceptual advantages. Finally, because most of the data here come from studies covering only one grade, we were not able to directly test assumptions about invariance across grades. This will be an important area for future work.

In sum, this study provides support for a theory of classroom interactions and has important implications for educational theory, research, and practice. With the growing focus on teacher effectiveness and accountability, it will be important for observational research on teaching to continue to contribute to the ongoing policy and practice debates.

Note

Authors Bridget Hamre and Robert Pianta are authors of the CLASS measure and own a company that conducts CLASS trainings. The authors gratefully acknowledge the National Institute for Child Health and Human Development for support of the Study of Early Child Care and Youth Development, the Interagency Consortium on School Readiness for support of the MyTeachingPartner study, the National Institute for Early Education Research (NIEER), W. Steven Barnett, the Pew Charitable Trusts, and the Foundation for Child Development for their support of the SWEEP Study, and the U.S. Department of Education for its support of the Multi-State Study of Pre-Kindergarten. Links to Learning is funded by the National Institute of Mental Health (PI: Atkins; R01 MH073749). The New York City Study of Social and Literacy Development is supported by grants from the U.S. Department of Education in collaboration with the Centers for Disease Control (R305L030003) and the William T. Grant Foundation (# 2618) to J. Lawrence Aber (PI), Jones and Brown (co-PIs), and by a grant from the William T. Grant Foundation (#7520) to Jones and Brown (co-PIs). The RULER intervention study is funded by the William T. Grant Foundation (#8364) to Brackett (PI), and Rivers & Salovey (co-PIs). However, the contents of this study do not necessarily represent the positions or policies of the funding agencies, and endorsement by these agencies should not be assumed. The authors are grateful for the help of the many children, parents, teachers, administrators, and field staff who participated in these studies. Please address questions or correspondence to Bridget K. Hamre, University of Virginia, Center for Advanced Study of Teaching and Learning, 350 Old Ivy Way Suite 100, Charlottesville, VA 22903; phone: 434–982–4855; fax: 434–243–5480; email: hamre@virginia.edu.

References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, **333**(6045), 1034–1037.

Bell, C. A., Gitomer, D., McCaffrey, D., Hamre, B., & Pianta, B. (2011). *An argument approach to observation protocols*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, **57**(2), 111–127.

Bollen, K. A. (1990). *Structural equations with latent variables*. New York: Wiley.

Bowlby, J. (1969). *Attachment and loss: vol. **1**. Attachment.* New York: Basic Books.

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2010, January). *Using emotional literacy to improve classroom social-emotional processes*. Paper presented at the William T. Grant and Spencer Foundation Grantees Meeting, Washington, DC.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school, expanded edition.* Washington, DC: National Academies Press.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. **1**. Theoretical models of human development* (5th ed., pp. 993–1029). New York: Wiley.

Brophy, J. (1999). *Teaching.* Educational Practices Series 1. Paris: International Academy of Education and International Bureau of Education.

Brophy, J., & Evertson, C. (1976). *Learning from teaching: A developmental perspective*. Boston: Allyn & Bacon.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs program. *Journal of Educational Psychology*, **102**, 153–167.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, **79**(4), 474–482.

Buyse, E., Verschueren, K., Doumen, S., Van Damme, J., & Maes, F. (2008). Classroom problem behavior and teacher-child relationships in kindergarten: The moderating role of the classroom climate. *Journal of School Psychology*, **46**(4), 367–391.

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher-student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of School Psychology*, **48**, 457–482.

Cappella, E., Frazier, S., Atkins, M., Schoenwald, S., & Glisson, C. (2008). Enhancing schools' capacity to support children in poverty: An ecological model of school-based mental health services. *Administration and Policy in Mental Health and Mental Health Services Research*, **35**, 395–409.

Cappella, E., Hamre, B. K., Kim, H., Henry, D. B., Frazier, S. L., Atkins, M. S., & Schoenwald, S. K. (2012). Teacher consultation and coaching within mental health practice: Classroom and child effects in urban elementary schools. *Journal of Consulting & Clinical Psychology*, **80**(4), 597–610.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (2001). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, **3**, 331–361.

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, **80**(1), 219–251.

Clare, L., & Aschbacher, P. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment*, **7**(1), 39–59.

Cochran-Smith, M., & Zeichner, K. (Eds.). (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahwah, NJ: Erlbaum.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self process and development: The Minnesota symposia on child development* (Vol. **23**, pp. 43–77). Hillsdale, NJ: Erlbaum.

Connor, C. M., Morrison, F. J., Fishman, B., Ponitz, C. C., Glasney, S., Underwood, P., . . . Schatschneider, C. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, **38**(2), 85–99.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

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

Douglas, K. (2009). Sharpening our focus in measuring classroom instruction. *Educational Researcher*, **38**, 518–521.

Downer, J. T., Sabol, T., & Hamre, B. K. (2010). Teacher-child interactions in the classroom: Toward a theory of within- and cross-domain links to children's developmental outcomes. *Early Education and Development*, **21**, 699–723.

Early, D., Maxwell, K., Burchinal, M., Alva, S., Bender, R., Bryant, D., . . . Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, **78**(2), 558–580.

Eccles, J. S., & Roeser, R. W. (1999). School and community influences on human development. In M. H. Boorstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (4th ed., pp. 503–554). Hillsdale, NJ: Erlbaum

Emmer, E. T., & Strough, L. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, **36**(2),103–112.

Evertson, C., Emmer, E., Sanford, J., & Clements, B. (1983). Improving classroom management: An experiment in elementary classrooms. *Elementary School Journal*, **84**, 173–188.

Gazelle, H. (2006). Class climate moderates peer relations and emotional adjustment in children with an early childhood history of anxious solitude: A child-by-environment model. *Developmental Psychology*, **42**, 1179–1192.

Good, T. L., & Brophy, J. E. (2008). *Looking in classrooms* (10th ed.). Boston: Pearson Allyn & Bacon.

Grossman, P. (2011). Framework for teaching practice: A brief history of an idea. *Teachers College Record*, **13**(12), 2836–2843.

Grossman, P., Loeb, S., Cohen, J., Hammerness, K., Wyckoff, J., Boyd, D., & Lankford, H. (2010). *Measure for measure: The relationship between measures of instructional practice in middle school English Language Arts and teachers' value-added scores.* Retrieved from http://www.caldercenter.org/upload/CALDERWorkPaper_45.pdf

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Development*, **76**(5), 949–967.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. Pianta, M. Cox, & K. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49–84). Baltimore: Brookes.

Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., Locasale-Crouch, J. L., Downer, J. T., . . . Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. *American Education Research Journal*, **49**(1), 88–123.

Hanushek, E. A. (2002). *The long-run importance of school quality.* NBER Working Paper 9071, National Bureau of Economic Research.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, **26**(4), 430–511.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, **6**, 53–60.

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

Jones, S. M., Brown J. L., & Aber, J. L. (2011). Two year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, **82**(2), 533–554.

Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teachers: Combining high-quality observations with student surveys and achievement gains.* Policy and practice brief prepared for the Bill and Melinda Gates Foundation.

Kaur, B. (in press). Equity and social justice in teaching and teacher education. *Teaching and Teacher Education.*

Kulik, J. A., & Kulik, C.-L. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, **58**(1), 79–97.

LaParo, K. M., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., Bryant, D. Early, D., . . . Burchinal, M. (2009). Quality in kindergarten classrooms: Observational evidence for the need to increase children's learning opportunities in early education classrooms. *Early Education and Development*, **20**, 657–692.

Li-Grining, C., Raver, C. C., Champion, K., Sardin, L., Metzger, M. W., & Jones, S. M. (2010). Understanding and improving classroom emotional climate in the "real world": The role of teachers' psychosocial stressors. *Early Education and Development*, **21**, 65–94.

Malmberg, L.-E., & Hagger, H. (2009). Changes in student teachers' agency beliefs during a teacher education year, and relationships with observed classroom quality, and day-to-day experiences. *British Journal of Educational Psychology*, **79**, 677–694.

Malmberg, L.-E., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2011). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology*, **102**, 916–932.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, **79**(3), 732–749.

Matsumura, L. C., Garnier, H., Slater, S. C., & Boston, M. B. (2008). Measuring instructional interactions at-scale. *Educational Assessment*, **13**(4), 267–300.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice*, **41**, 226–233.

NICHD Early Child Care Research Network. (2002). The relation of first-grade classroom environment to structural classroom features, teacher, and student behaviors. *Elementary School Journal*, **102**, 367–387.

NICHD Early Child Care Research Network. (2003). Social functioning in first grade: Prediction from home, child care, and concurrent school experience. *Child Development*, **74**, 1639–1662.

NICHD Early Child Care Research Network. (2005). A day in third grade: A large-scale study of classroom quality and teacher and student behavior. *Elementary School Journal*, **105**, 305–323.

Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J. (2010). A validation of the classroom assessment scoring system in Finnish kindergartens. *Early Education and Development*, **21**, 95–124.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, **74**, 525–546.

Pianta, R. C., Belsky, J., Houts, R., Morrison, F., & NICHD Early Child Care Research Network. (2007). Opportunities to learn in America's elementary classrooms. *Science*, **315**, 1795–1796.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, **38**, 109–119.

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

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system (CLASS)*. Baltimore: Brookes.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, **23**, 431–451.

Ponitz, C. C., Rimm-Kaufman, S. E., Brock, L. L., & Nathanson, L. (2009). Contributions of gender, early school adjustment, and classroom organizational climate to first grade outcomes. *Elementary School Journal*, **110**, 143–162.

Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review*, **38**, 102–120.

Pressley, M., Roehrig, A., Raphael, L., Dolezal, S., Bohn, C., Mohan, L., . . . Hogan, K. (2003). Teaching processes in elementary and secondary education. In W. Reynolds & G. Miller (Eds.), *Handbook of psychology: Volume 7. Educational psychology* (pp. 153–176). Hoboken, NJ: Wiley.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Smallwood, K., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, **23**(1), 10–26.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*, **77**(2), 302–316.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, **104**(3), 700–712.

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, **45**, 958–972.

Rimm-Kaufman, S. E., & Hamre, B. K. (2010). The role of psychological and developmental science in efforts to improve teacher quality. *Teachers College Record*, **112**(12), 2988–3023.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Rowan, B., & Correnti, R. (2009). Measuring reading instruction with teacher logs. *Educational Researcher*, **38**(7), 549–551.

Rudasill, K. M., Gallagher, K., & White, J. M. (2010). Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *Journal of School Psychology*, **48**, 113–134.

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

Skibbe, L., Behnke, M., & Justice, L. M. (2004). Parental scaffolding of children's phonological awareness skills: Interactions between mothers and their preschoolers with language difficulties. *Communication Disorders Quarterly*, **25**(4), 189–203.

Strong, M., Gargani, J., & Hacifazlioğlu, Ö. (2011). Do we know a successful teacher when we see one? Experiments in the identification of effective teachers. *Journal of Teacher Education*, **62**(4), 367–382.

Stuhlman, M., Downer, J., & DeCoster, J. (2012, April). *Developing and validating scalable approaches to assessing effective classroom interactions*. Paper presented at the meeting of the American Educational Research Association, Vancouver, BC.

Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal*, **104**, 3–28.