

2014-12-15

Examining the Association Between Variability in Classroom Interaction Quality and Problem Behavior in Head Start Children

Veronica A. Fernandez

University of Miami, vfernandez17@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation

Fernandez, Veronica A., "Examining the Association Between Variability in Classroom Interaction Quality and Problem Behavior in Head Start Children" (2014). *Open Access Dissertations*. 1342.

https://scholarlyrepository.miami.edu/oa_dissertations/1342

This Embargoed is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.

UNIVERSITY OF MIAMI

EXAMINING THE ASSOCIATION BETWEEN VARIABILITY IN CLASSROOM
INTERACTION QUALITY AND PROBLEM BEHAVIOR IN HEAD START
CHILDREN

By

Veronica A. Fernandez

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

December 2014

©2014
Veronica A. Fernandez
All Rights Reserved

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

EXAMINING THE ASSOCIATION BETWEEN VARIABILITY IN CLASSROOM
INTERACTION QUALITY AND PROBLEM BEHAVIOR IN HEAD START
CHILDREN

Veronica A. Fernandez

Approved:

Rebecca J. Bulotsky-Shearer, Ph.D.
Associate Professor of Psychology

Daryl B. Greenfield, Ph.D.
Professor of Psychology

Maria M. Llabre, Ph.D.
Professor of Psychology

Michael Alessandri, Ph.D.
Clinical Professor of Psychology

Batya E. Elbaum, Ph.D.
Professor of Teaching and Learning

M. Brian Blake, Ph.D.
Dean of the Graduate School

FERNANDEZ, VERONICA A.

(Ph.D., Psychology)

Examining the Association Between Variability
in Classroom Interaction Quality and Problem
Behavior in Head Start Children

(December 2014)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Associate Professor Rebecca J. Bulotsky-Shearer.

No. of pages in text. (60)

The present study examined the role of variability in observed classroom interaction quality (as assessed by the CLASS) in a sample of Head Start children ($N = 965$ children, within 54 classrooms and across 8 centers). First, the psychometric properties of the typically used classroom quality mean scores were examined to test whether the levels (mean differences) and rank ordering (reliability) of the scores varied across activity settings (Free Play, Whole Group, Small Group, Meals or Routines/Transitions). Then, variability scores were calculated for each of the three classroom quality domains to examine the unique association between variability in classroom interaction quality and behavior problems, above and beyond mean classroom interaction quality. For classroom interaction quality means, results indicate that across activity settings the scores were reliable for all three CLASS domains, but the mean levels varied for two domains: Emotional Support and Classroom Organization. Results from the multilevel analyses suggest that variability in classroom interaction quality differentially related to children's social-emotional outcomes, above and beyond mean levels. Findings provide additional evidence for the reliability of the CLASS and contribute to our understanding of the important role of variability in classroom interaction quality for children with behavior problems in Head Start.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	iv
Chapter	
1 INTRODUCTION.....	1
Quality in the Preschool Classroom	2
Theoretical Framework.....	4
Role of Classroom Interaction Quality.....	6
Preschool Behavior Problems.....	10
Variability in Classroom Interaction Quality.....	13
Present Study.....	15
2 METHOD	17
Procedure.....	17
Participants	17
Measures.....	18
Data Analytic Plan.....	20
3 RESULTS.....	25
Descriptive Statistics	25
Psychometric Properties of Mean Classroom Interaction Quality Scores.....	25
Unique Associations between Variability in Classroom Interaction Quality and Behavior Problems.....	26
4 DISCUSSION.....	31
Level of Classroom Interaction Quality Across Activity Settings	31
Reliability of Mean Classroom Interaction Quality Across Activity Settings	35
Association Between Classroom Interaction Quality Behavior Problems	36
Limitations and Future Directions.....	41
Implications for Policy and Practice.....	43
REFERENCES.....	47

LIST OF TABLES

Table	Page
1 DESCRIPTIVE STATISTICS.....	55
2 ACTIVITY SETTING MEANS.....	56
3 GENERALIZABILITY THEORY RESULTS.....	57
4 BIVARIATE CORRELATIONS.....	58
5 MULTILEVEL MODELS FOR DISRUPTIVE BEHAVIOR.....	59
6 MULTILEVEL MODELS FOR DISCONNECTED BEHAVIOR.....	60

Chapter 1: Introduction

National studies suggest that many children entering kindergarten are not yet “ready to learn” because they lack the necessary social-emotional skills to participate successfully in classroom learning contexts (e.g., Raver & Knitzer, 2002; Rimm-Kaufman, Pianta, Cox, 2000). Difficulties with social-emotional adjustment in preschool, such as behavior problems, may interfere with children’s engagement in social interactions and learning within the classroom (Gilliam, 2005; Thompson & Raikes, 2007). As a consequence, children displaying early behavior problems are at risk for concurrent and long-term negative social and academic outcomes (e.g. Denham, 2006; Lutz, Fantuzzo, McDermott, 2002; Raver, 2002). This concern is heightened within preschool programs serving children from economically disadvantaged backgrounds, where up to 34% percent of children exhibit behavior problems as a function of their disproportionate exposure to ecological risks associated with living in poverty (Barbarin, 2007; Campbell, Shaw, & Gilliom, 2000; Qi & Kaiser, 2003; Webster-Stratton & Hammond, 1998).

Fortunately, early intervention research suggests that early childhood teachers can play a critical role in reducing classroom behavior problems by providing consistency and predictability within the classroom, especially for children facing the greatest level of behavioral risks (e.g., Powell, Dunlap & Fox, 2006). Research suggests that greater levels of consistency are associated with increased social competence and decreased behavior problems (Fox, Dunlap, Hemmeter, Joseph & Strain, 2003; Johnson, Stoner & Green, 1996; Kern & Clemens, 2007). Best practices in early childhood education also align with this research and recommend that children experience consistency and

predictability within high quality classroom environments (National Association for the Education of Young Children, 2009). Yet preschool classrooms are complex and dynamic environments. A typical preschool day is comprised of various activity settings (e.g., whole group, small group, free play) that children transition through, each with unique demands on teachers and children (Howes & Smith, 1995; Kontos & Wilcox-Herzog, 1997; Vitiello, Booren, Downer, Williford, 2012). The level of classroom quality that children experience may fluctuate as a function of the structure of the typical preschool routine. Therefore, rather than experiencing consistency over the course of the day, children may experience variability in the level of classroom quality. However, research examining how such variability in classroom quality might influence children's social-emotional outcomes is very limited, particularly for children from low-income households. The purpose of the present study was to examine the role of variability in observed classroom quality (as measured by the Classroom Assessment Scoring System; Pianta, La Paro, & Hamre, 2008) and its association with behavior problems, in a sample of linguistically and culturally diverse preschool children from low-income backgrounds.

Quality in the Preschool Classroom

Preschool classroom quality is generally understood as comprising two broad categories: *structural* features of quality and *process* features of quality. Structural quality includes aspects of a classroom or program that are typically regulated by state agencies. Some examples of structural features of quality include teacher-child ratios, group sizes, teacher credentials, the use of a curriculum, and availability of play and instructional materials (Harms, Clifford, & Cryer, 1998). Although structural features are important, there has been an increased national focus on the importance of classroom

process features, or the quality of the interactions that occur within the classroom that support children's learning. Process features of quality mainly focus on the dynamic teacher-child interactions within the classroom (Hamre & Pianta, 2007). In classrooms with high interaction quality, teachers are warm, sensitive and responsive, provide clear expectations and predictable routines and facilitate cognitively stimulating and engaging learning activities (Howes et al., 2008; Mashburn, 2008; Mashburn & Pianta, 2006). National research studies demonstrate that classroom interaction quality is important because it directly influences children's social-emotional and academic learning, mediating the relationship between structural features of classroom quality and children's outcomes (Pianta et al., 2005). In other words, although structural quality is important, it is through the quality of the instructional and emotionally supportive interactions within the classroom that children derive the greatest academic and social-emotional benefits.

Given that young children learn most directly through social interactions in their immediate environment (Bronfenbrenner & Morris, 1998; Vygotsky, 1978) and considering the growing number of young children enrolled in preschool (National Institute for Early Education Research [NIEER], 2011), classrooms are ideal contexts for intervention. Head Start, the nation's largest and most comprehensive early intervention program for children from low-income backgrounds plays a critical role as one of the earliest mechanisms to promote children's social-emotional adjustment within the preschool classroom context (U. S. Department of Health and Human Services, 2002). In fact, the Head Start framework embodies a whole child approach, focusing on all developmental domains, including social and emotional development (National Education Goals Panel, 1997).

Within the context of Head Start classrooms, research highlights the important role of high quality interactions between teachers and children in fostering school readiness. It is well documented that high classroom interaction quality is positively associated with social competence and negatively associated with behavior problems (e.g., Downer, Sabol and Hamre, 2010; Howes et al., 2011; Pianta, 1999). In particular, differential benefits are afforded to children at risk; findings suggest that these associations are stronger for children from low-income backgrounds, indicating that high quality classroom interactions are especially critical for the population of children served by Head Start (Burchinal et al., 2008; Howes et al., 2008).

Theoretical Framework

The bioecological model and attachment theory provide an ideal framework for understanding the critical role of classroom interaction quality on children's social-emotional adjustment. According to the bioecological model, children's learning experiences are embedded in a larger ecological context, consisting of proximal and more distal systems. The processes which most directly influence children are those occurring within proximal contexts, such as the home or preschool classroom (Bronfenbrenner & Hedges, 2004; Bronfenbrenner & Morris, 2006). Within the preschool classroom, dynamic interactions with teachers have been identified as the primary mechanism for development and learning (Mashburn, 2008; Pianta, 1999). For example, high quality interactions are associated with favorable social and academic outcomes (e.g., Downer et al., 2010; Howes et al., 2011). Children in classrooms with high classroom interaction quality experience warm, sensitive, responsive and cognitively stimulating interactions,

within the context of predictable routines and engaging learning activities. These high quality interactions with teachers facilitate children's engagement in social interactions and learning within the classroom.

Attachment theory underscores the role of consistency in relationships between children and their primary caregivers (Ainsworth, 1969; Bowlby, 1969). Typically this pertains to parent/guardian-child relationships, but as children increasingly spend more time in preschool classrooms (NIEER, 2011), attachment theory can be extended to the teacher-child relationship as well. In fact, Ainslie (1990) found that children were as likely to develop secure attachments to child care providers (such as teachers) as they were to parents. According to the attachment theory, children develop secure attachment as a result of social interactions with sensitive, responsive and consistent caregivers (Ainsworth, Belhar, Waters & Wall, 1978; Bowlby, 1969). When young children's needs are consistently met, they feel secure about the dependability of their caregiver, establishing a "secure base" that enables them to feel comfortable to explore and learn in their environment. The quality of these early relationships has an important influence on children's social-emotional and academic development. Research suggests that secure attachment is positively associated with social competence and negatively associated with ratings of behavior problems (e.g., Cutrona, Colangelo, Assouline, & Russell, 1994). With respect to the preschool classroom, children who do not experience consistency in the quality of interactions with teachers may then not experience the social-emotional benefits that are associated with predictable, sensitive interactions. This makes explicit

the need to examine empirically the association between the potential lack of such consistency, or variability, in classroom interaction quality and children's social-emotional outcomes.

The bioecological model and attachment theory support the notion that the process through which children's academic and social development is fostered is dynamic and bidirectional, including both the child and the adult, who are contributing to these interactions. For children exhibiting behavior problems, teachers may play an even more important role in individualizing support and instruction. Research findings suggest that teachers' decisions about how they interact with these children, structure the daily routine and plan activities influence children's ability to successfully engage in social interactions and learning within the classroom (Powell, Burchinal, File, & Kontos, 2008; Vitiello et al., 2012). For children with behavior problems, experiencing consistency in classroom interactions and in turn, knowing what to expect within their environment, may support more adaptive behavior and success in negotiating the demands of classroom. Given that behavior problems may interfere with children's ability to benefit from naturally occurring learning experiences within the preschool classroom, the teacher's role in providing a stable and supportive learning environment may be more critical for children displaying behavior problems. However, little is known about the association between variability in classroom interaction quality within Head Start and children's behavior problems.

Role of Classroom Interaction Quality

Associations with social-emotional outcomes. Research suggests that high quality classroom interactions within the preschool classroom are associated with social-

emotional adjustment (e.g, Burchinal et al., 2008; Curby et al., 2009; Gazelle, 2006; Mashburn et al., 2008; NICHD ECCRN, 2003; Wilson, Pianta, & Stuhlman, 2007). For example, Howes et al. (2008) found that preschool children exhibited lower aggressive behaviors with peers and higher social competence when teachers reported positive teacher-child relationships with them. O'Connor, Dearing, and Collins (2011) in their longitudinal study also found that teacher-reported positive teacher-child relationships in preschool were associated with fewer early behavior problems; in addition positive teacher-child relationships decreased the likelihood of children with high levels of behavior problems in preschool from developing trajectories of long-term behavior problems in elementary school.

Other findings from early childhood research also suggest that preschool children in classrooms with high observed levels of classroom interaction quality (as measured by the CLASS) exhibit higher self-regulation, increased behavioral control (Rimm-Kaufman, Curby, Grimm, Nathanson & Brock, 2009), decreased problem behavior (e.g., Howes et al., 2008; Mashburn et al., 2008) and engage in more advanced social conversations (Rimm-Kaufman, LaParo, Downer, & Pianta, 2005). Together these research studies, using multiple methods and sources to measure classroom interaction quality, suggest that high quality interactions with teachers are associated with more favorable social-emotional adjustment during early childhood.

Head Start children. High classroom interaction quality during preschool is particularly important for at-risk children (e.g., Burchinal et al., 2008; Dickinson & Tabors, 2001; Hamre & Pianta, 2005; Howes et al., 2008; Stipek et al. 1998). Although research indicates that high classroom interaction quality is positively associated with

social-emotional development in general, there is evidence to suggest that high classroom interaction quality is more important for children from low-income backgrounds (Burchinal et al., 2008; Howes et al., 2008; Lamb & Ahnert, 2006; Pleuss & Belsky, 2009). In a first grade sample, Hamre and Pianta (2005) found that the association between high classroom interaction quality and child outcomes was greatest for children at higher levels of economic and academic risk. At-risk children (i.e., with low levels of maternal education and non-intact family structure) in classrooms with high to moderate levels classroom interaction quality performed as well as children not classified as at-risk (Hamre & Pianta, 2005). These findings suggest that high classroom interaction quality has the potential to compensate for the negative risks associated with poverty. Given the potential benefit to at-risk children during the preschool developmental period, it is critical to focus research and intervention efforts on examining and improving classroom interaction quality for children from economically disadvantaged backgrounds, such as those children participating in Head Start programs.

Observational measure. The Classroom Assessment Scoring System (CLASS), a widely used observational measure, assesses three broad domains: Emotional Support (or Emotional Climate), Classroom Organization and Instructional Support (Pianta et al., 2008), as key indicators of classroom interaction quality (La Paro, Pianta, & Stuhlman, 2004). Classrooms with high Emotional Support are characterized by warm, sensitive and responsive classroom interactions. Classroom Organization assesses how well teachers structure and manage classroom activities, routines, and materials to support children's self-regulation, attention, and engagement in learning throughout the day. Classrooms with high Instructional Support are characterized by interactions where teachers stimulate

higher-order thinking and problem-solving, provide appropriate feedback, engage in in-depth discussions, make learning meaningful to children and model the use of advanced language. The three CLASS domains of classroom interaction quality have been found to differentially relate to preschool children's academic and social development (e.g., Downer, et al., 2010).

CLASS domains and social-emotional outcomes. Research provides evidence for within domain specificity in the associations between the three broad CLASS domains and social and academic outcomes during early childhood. Stronger differential associations with child outcomes have been found when the components of quality that are measured more closely align with the developmental domain examined (Downer, et al., 2010). For example, there are stronger associations between Emotional Support and social-emotional outcomes, as compared to Instructional Support and social-emotional outcomes. In addition, stronger associations have been found when Instructional Support is used to predict academic outcomes, than when Emotional Support is used (Downer, et al., 2010).

Accordingly, higher Emotional Support has been consistently associated with higher social competence (e.g, Burchinal et al., 2008; Curby et al., 2009; Wilson, Pianta, & Stuhlman, 2007) and lower behavior problems (e.g., Gazelle, 2006; Mashburn et al., 2008; National Institute of Child Health and Human Development Early Child Care Research Network, 2003). In addition, positive associations have been found between Classroom Organization and social outcomes. Classroom Organization has been associated with higher self-regulation and increased behavioral control (Rimm-Kaufman, et al., 2009). Additionally, components of Classroom Organization (e.g., clarity of rules,

proactive behavior management, and predictable routines), are considered to play an important role in supporting the development of prosocial behavior and the formation of positive social relationships (Donohue, Perry & Weinstein, 2003; Raver, & Zigler, 1997; Webster-Stratton et al., 2001). On the other hand, findings for associations between Instructional Support and social outcomes are mixed. A few studies have found higher Instructional Support to be associated with higher social competence (e.g., Guthrie et al., 2000; Pianta, La Paro, Payne, Cox, & Bradley, 2002), although findings from other studies do not confirm such an association (e.g., Mashburn et. al. 2008).

Preschool Behavior Problems

Early social-emotional adjustment difficulties, such as behavior problems, can interfere with children's engagement in social interactions and learning within the preschool classroom (Gilliam, 2005; Rimm-Kaufman, et al., 2000). Negative associations between behavior problems and children's social-emotional and academic outcomes have been consistently documented in samples from low-income backgrounds (e.g., Bulotsky-Shearer, Fernandez, Dominguez, & Rouse, 2011; Dobbs, Doctoroff, Fisher, & Arnold, 2006; Fantuzzo, Bulotsky, McDermott, Mosca, & Lutz, 2003; Lonigan, et al., 1999; McWayne & Cheung, 2009). For example, in recent Head Start studies, preschool behavior problems have been negatively associated with peer social competence (e.g., Bulotsky-Shearer, Fantuzzo, & McDermott 2008), literacy and mathematics skills, as well as approaches to learning (e.g., Bulotsky-Shearer, Fernandez, Domínguez, & Rouse, 2011; Domínguez, Vitiello, Fuccillo, Greenfield, & Bulotsky-Shearer, 2011). Without early identification and intervention, behavior problems tend to be stable over time (e.g., Achenbach, Edelbrock, & Howell, 1987; Campbell & Ewing, 1990; Richman, Stevenson,

& Graham, 1982) and are associated with long-term social-emotional and academic maladjustment (Bulotsky-Shearer & Fantuzzo, 2011; Campbell & Ewing, 1990; De Feyter, 2011).

Externalizing and internalizing behavior problems. Early childhood research makes a distinction between two broad types of behavior problems: externalizing and internalizing. Externalizing behavior problems typically pertain to outward acts of aggression, disruption, tantrums, and over activity (e.g., Achenbach & Edelbrock, 1992). Internalizing behavior problems are typically defined by shyness, flat affect, anxiety and social withdrawal (e.g., Achenbach & Edelbrock, 1981; Hinshaw, Han, Erhardt & Huber, 1992).

Research within Head Start has identified differential associations between the two types of behavior problems and children's social-emotional and academic outcomes. For example, research suggests that Head Start children with externalizing behavior problems in the fall of the preschool year exhibit lower attention, persistence and language ability in the spring (Fantuzzo, Bulotsky-Shearer, Fusco, & McWayne, 2005). Whereas, children with internalizing behavior problems early in the preschool year have been found to exhibit lower competence motivation, affective engagement, adaptive emotional regulation, literacy and mathematics skills at the end of the year (Domínguez, Vitiello, Maier, & Greenfield, 2010; Fantuzzo et al., 2003, 2007). In preschool settings, behavior problems can be readily observed and assessed within the context of peer play. Children with externalizing behavior problems tend to engage in *disruptive* play with peers and children with internalizing behavior problems tend to display *disconnected* behaviors within the context of peer play (Fantuzzo et al., 2005).

The role of consistency for children exhibiting behavior problems. Early childhood researchers document the importance of consistency in the classroom environment, particularly for children exhibiting behavior problems (e.g., Powell, Dunlap & Fox, 2006; Kern & Clemens, 2007). Greater levels of consistency are associated with increased social competence and decreased behavior problems (Fox, Dunlap, Hemmeter, Joseph & Strain, 2003; Johnson, Stoner & Green, 1996; Kern & Clemens, 2007). In other words, when children are provided with predictable routines and have clear rules and consequences for their behavior, they understand what is expected of them, are better able to anticipate what may happen next, and therefore may be better able to regulate their behavior. Based on this research, providing children with consistent routines and expectations within the classroom may facilitate more opportunities for children's to be actively and positively engaged in classroom learning activities and social interactions, and may reduce the likelihood that children will display behavior problems.

In summary, evidence suggests that behavior problems may interfere with children's ability to benefit from naturally occurring learning opportunities within the preschool classroom. Given that young children learn most directly through their immediate environment, behavior problems in preschool place may children at risk for long-term social and academic consequences, particularly for children from low-income backgrounds. Fortunately, research also indicates that high quality early educational experiences, such as participation in Head Start, may serve as a buffer and protective influence to support more positive outcomes for children at risk (e.g., Downer, Rimm-Kaufman, & Pianta, 2007; Hamre & Pianta, 2005; Rimm-Kaufman et al., 2002).

However, to date, research has not examined the extent to which variability in the quality of interactions that children in Head Start classrooms may naturally experience over the course of a typical day may contribute to behavior problems.

Variability in Classroom Interaction Quality

In research examining associations between classroom interaction quality and children's outcomes using the CLASS, an overall mean score typically is calculated for each of the three classroom quality domains: Emotional Support, Classroom Organization, and Instructional Support (as delineated in the CLASS manual). The mean score for each domain is created by averaging each of four scores based on four sequential observation cycles (Pianta et al., 2008). However, this average score, reflecting the overall quality of classroom interactions, may not capture the fluctuations (or variability) in quality that children likely experience throughout the observation period (usually across a 3-4 hour period during the morning). Although it is well documented that high overall mean levels of classroom interaction quality are positively associated with children's social-emotional development (e.g., Burchinal et al., 2008; Curby et al., 2009; Gazelle, 2006; Mashburn et al., 2008), very little is known about whether variability in classroom interaction quality meaningfully relates to children's social-emotional outcomes.

Because preschool classrooms are complex and dynamic environments, there is a need to look beyond mean levels of classroom interaction quality and to examine whether variability in classroom interaction quality across a typical morning plays a role in children's social-emotional development. To date, there is only one preschool study that has examined this empirically and this study focused on the Emotional Support domain

of the CLASS only. Curby et al. (2013) calculated what they termed “within-day consistency” for the Emotional Support domain of the CLASS. In this study, within-day consistency was conceptualized as the degree to which preschool classroom interactions within the Emotional Support domain were rated at the same level of quality across the four observation cycles across a typical morning. “Within-day consistency” scores were created by calculating the variance across scores from four sequential observation cycles and then reversing the valence. This was not meant to measure whether teachers were engaging in the same types of interactions or the same activities, but rather to capture the consistency in the quality of the classroom interaction across the observation cycles.

With respect to children’s outcomes, Curby et al. (2013) found that within-day consistency in Emotional Support was positively associated with both social and academic outcomes. Higher consistency in emotionally supportive interactions were associated with higher teacher-reported social competence (in kindergarten) as well as direct assessments of expressive language and early literacy (letter naming and rhyming) outcomes (in preschool), above and beyond mean levels of Emotional Support.

Findings from Curby et al.’s (2013) initial study suggest that within early childhood classrooms it may be important to examine variability in classroom interaction quality as it relates to children’s outcomes. However, this research has several limitations with respect to preschool children enrolled in Head Start exhibiting behavior problems. First, the potential role of activity settings within the classroom on the overall classroom quality that children experience was not examined. Second, the study focused only on the Emotional Support CLASS domain. It is important to examine whether variability in the other CLASS domains (e.g., Classroom Organization and Instructional

Support) are associated with children's developmental outcomes. In addition, Curby et al. (2013) only examined the association between consistency in preschool classroom interaction quality and social-emotional outcomes (social competence and behavior problems) in kindergarten. It is critical to examine the associations between variability in preschool classroom interaction quality and children's social-emotional outcomes during the preschool year. Finally, although Curby et al. (2013) examined these associations for a large sample of preschool children participating in the Multi-State Study of Pre-Kindergarten and the State-Wide Early Education Programs Study, only a small percentage of the children participating were enrolled in Head Start or from culturally, or linguistically diverse low-income backgrounds. Given that these children are at increased risk for exhibiting classroom behavior problems, research is needed that includes a more diverse sample of children from low-income backgrounds.

Present Study

To address this need, the purpose of the present study was to extend prior research by examining the role of variability in classroom interaction quality and preschool behavior problems, in a sample of linguistically and culturally diverse children from low-income backgrounds. All three classroom interaction quality domains were used, including Emotional Support, Classroom Organization and Instructional Support. The role of variability was examined in two ways. First, the psychometric properties of the typically used classroom interaction quality mean scores were examined to answer the following two research questions: 1) does mean classroom interaction quality vary as a function of activity setting? and 2) are mean classroom interaction quality scores reliable across the various activity settings? Once the psychometric properties of the mean

classroom interaction quality domain scores were examined, a variability score was calculated for each of the three classroom interaction quality domains to address the third research question: 3) what is the unique association between variability in classroom interaction quality and behavior problems, above and beyond mean classroom interaction quality? Based on previous research and theory, it was expected that classroom interaction quality would vary as a function of activity setting and that means would not be reliable across activity settings (e.g., Vitiello et al., 2012). Finally, higher variability in classroom interaction quality was expected to be associated with higher behavior problems at the end of the year (e.g., Curby et al., 2013).

Chapter 2: Method

Procedure

This research project was part of a larger University-Head Start collaborative research project in Miami-Dade County. Approval for this project was obtained from the University's Institutional Review Board (IRB), from the Director of the local Head Start Program, and from the Head Start Program's Parent Policy Council. In the early fall of the preschool year, members of the research team met with the directors of each Head Start center to obtain consent for participation in the research study. The research team then explained the study and obtained consent for those teachers who were willing to participate. Parental consent was then obtained for children with the assistance of teachers. In the fall and spring, teachers completed ratings of children's behavior. In the winter, trained observers conducted observations of classroom interaction quality.

Participants

Participants included a representative sample of children from a culturally and linguistically diverse, urban Head Start program in the Southeastern United States ($N = 965$ children, across 54 classrooms and 8 centers). Approximately 50% of the children were female and children's mean age in the fall was 48.07 months ($SD = 6.86$). Children were predominantly Black and Hispanic (45% and 44%, respectively), with 6 % White non-Hispanic and 4% identified as being of another ethnicity (including multi-racial, Asian, Other, or Native Islander). Approximately 40% percent of the children spoke Spanish at home. All children were from families that met the federal poverty criteria for enrollment in the Head Start program.

Measures

Classroom-level problem behavior. In the fall and spring, teachers completed the Penn Interactive Peer Play Scale, Teacher version (PIPPS-T; Fantuzzo, Coolahan, Mendez, McDermott, & Sutton-Smith, 1998). The PIPPS-T is a 32-item rating scale used to assess children's peer interactive behaviors within the classroom context, with three scales: Disruption and Disconnection, and Interaction (each demonstrating high internal consistency for use within Head Start samples; Cronbach's alpha = .92, .91, and .89, respectively). Because the focus of the present study was to examine problem behavior, only the Disruption and Disconnection scales were used (the Interaction scale measures prosocial behaviors). Items on the Disruption scale reflect externalizing problem behavior such as "starts fights and arguments" and "disrupts play of others." Items on the Disconnection scale reflect internalizing problem behavior such as "hovers outside play group," and "wanders aimlessly." For the present study, *T* scores were calculated (based on the standardization sample; Fantuzzo et al., 1998) and used in the analysis. In addition, individual child-level *T* scores were aggregated at the classroom level; both child- and classroom-level problem behavior scores were included in the final models.

Classroom interaction quality. In the winter, the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) was used to assess the quality of the interactions between teachers and children. The CLASS is a standardized observational measure of three domains of classroom interaction quality: (1) Emotional Support, consisting of four dimensions: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspective; (2) Classroom Organization, consisting of three dimensions:

Behavior Management, Productivity, and Instructional Learning Formats; and (3) Instructional Support, consisting of three dimensions: Concept Development, Quality of Feedback and Language Modeling (scores demonstrate acceptable internal consistency of .89, .77, .83, respectively; La Paro et al., 2004). There is strong evidence for reliability and validity in Head Start samples (Pianta et al., 2005) and in bilingual, Spanish-speaking preschool classrooms (Downer, López, Grimm, Hamagami, Pianta, & Howes, 2012).

Observations were conducted as delineated in the CLASS manual. For each classroom, four, sequential 30-minute cycles (20 minutes of observing and 10 minutes of coding) were conducted by certified observers across one morning. Prior to conducting classroom observations, research assistants completed a two-day intensive training conducted by a CLASS-certified trainer at the University. Observers were “certified” by successfully achieving at least 80% reliability with the Master Codes on a set of five online videos. To minimize observer drift across the course of the observation period and to ensure at least 80% interrater reliability, 20% of classrooms were double coded. Each dimension was rated on a 7-point scale: low quality (1, 2), mid-range (3–5), or high quality (6, 7). In addition to the quality ratings, observers indicated the activity setting (Free Play, Whole Group, Small Group, Meals or Routines/Transitions) at the start of the observation cycle. Subsequent activity settings were also noted, including the time spent in and transitioning between each setting. At the end of each cycle, the activity setting with the greatest amount of time was designated as the primary activity setting.

CLASS mean scores. Means scores for each of the three CLASS domains were calculated for each classroom as delineated in the CLASS manual. Specifically, the raw scores from each dimension comprising each domain were averaged across the four

observation cycles to obtain an overall mean (representing an average quality score in Emotional Support, Classroom Organization, and Instructional Support) as rated across a typical morning.

CLASS variability scores. Variability scores for each of the three CLASS domains were obtained by calculating the variance for each domain, using the raw scores for all four cycles, for each of the domain's respective dimensions. For example, the Classroom Organization domain has three dimensions: behavioral management, instructional learning formats, and productivity. Each of the three dimensions received a score (ranging from 1 to 7) for each of the 4 observation cycles. In total for Classroom Organization, each classroom received 12 individual scores (4 for Behavioral Management, 4 for Productivity, and 4 for Instructional Learning Formats) and all of these 12 scores were used to calculate the variance. The same procedure was followed to calculate the variability scores for the Emotional Support domain (4 cycles of scores across 4 dimensions, Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspective) and for the Instructional Support domain (4 cycles of scores across 3 dimensions, Concept Development, Quality of Feedback and Language Modeling).

Data Analytic Plan

Psychometric properties of mean classroom interaction quality scores. For the first research question, a series of analysis of variance (ANOVA) tests were conducted, examining whether the mean classroom interaction quality varied as a function of primary activity setting. Separate analyses were conducted for each CLASS domain (Emotional Support, Classroom Organization, and Instructional Support). In each model,

the mean scores across the six activity settings (free play, whole group, small group, meals, and routines/transitions) were compared. For each domain, the mean scores were the dependent variable and the activity settings were the levels for the fixed factor.

To address the second research question, generalizability theory (G theory; Cronbach, Gleser, Nanda, Rajaratnam, 1972) was employed to examine whether the obtained mean classroom interaction quality scores were reliable across activity settings. Adequate reliability coefficients would indicate that the rank ordering of the scores was consistent between activity settings. Again, separate analyses were conducted for each CLASS domain (Emotional Support, Classroom Organization, and Instructional Support).

In a classical measurement approach, an observed score is partitioned into variance attributed to individual differences and variance attributed to random measurement error (Novick, 1966). G theory provides an ideal framework to further partition the variance and examine whether it can be explained by specified sources of measurement error, or *facets*. For the present study, of interest was whether scores on the CLASS domains were reliable across one facet: activity settings. G theory analyses produces reliability estimates by examining the consistency of the rank ordering of scores across the conditions of specified facets and can also accommodate systematic differences across conditions. Here, the five types of activity settings (free play, whole group, small group, meals, and routines/transitions) were included in the analyses as the activity setting conditions.

The variance in classroom interaction quality (CLASS) means was partitioned into variance attributed to individual difference between classrooms (σ_c^2) and residual

variance ($\sigma_{a,ca,e}^2$), which included the variance attributed to activity setting (σ_a^2), the interaction between classroom (c) and activity setting (a), as well as random error (e) (Cronbach et al., 1972). Reliability coefficients are typically calculated for one instance of a variable: $\sigma_c^2 / (\sigma_c^2 + \sigma_{a,ca,e}^2)$. For the present study, four consecutive cycles of observations were conducted for each classroom (as delineated in the CLASS manual; Pianta et al., 2008). Therefore, reliability estimates were calculated to represent the reliability estimates for four cycles of observations: $\sigma_c^2 / (\sigma_c^2 + \sigma_{a,ca,e}^2 / 4)$.

Since four consecutive cycles of observations were conducted and an activity setting was designated for each of the four observation cycles, cycle and activity setting were confounded. Before conducting the analyses to examine the consistency of scores across the activity settings, it was necessary to test whether the rank ordering was consistent across cycle, or time. To do this, it was necessary to hold constant an activity setting, conduct a G theory analysis and calculate a reliability coefficient. The observed variance in CLASS means was partitioned into variance attributed to individual difference between classrooms (σ_c^2), variance attributed to time (σ_t^2) and residual variance ($\sigma_{ct,e}^2$), which included the interaction between classroom (c) and time (t), as well as random error (e). This analysis was repeated for each CLASS domain: Emotional Support; Classroom Organization; and Instructional Support. Once it was determined that CLASS means were sufficiently reliable across time (or cycle), it was appropriate to conduct the G theory analysis to examine whether scores on the CLASS domains were reliable across activity settings.

Unique association between variability in classroom interaction quality and behavior problems. To address the third research question, variability scores were first

calculated for each of the three classroom interaction quality domains, to examine the unique association between variability in classroom interaction quality (for each of the three CLASS domains: Emotional Support, Classroom Organization and Instructional Support) and behavior problems (for each of the two PIPPS-T scales: Disruption and Disconnection in the spring). Bivariate correlations were then used to examine the associations between the variability in the three CLASS domains of interaction quality and behavior problems. Finally, given the nested nature of the data (children within classrooms), multilevel modeling using the HLM Version 6.06 software (Raudenbush, Bryk, Cheong & Congdon, 2004) was employed to parse the variance in children's behavior problems into child-level and classroom-level variance (Heck, 2001).

Two-level models were estimated separately for the two outcomes (Disruption and Disconnection). For each outcome model, the three domains of teacher-child interaction quality (Emotional Support, Classroom Organization, and Instructional Support) were examined separately. First, unconditional models were estimated to determine the distribution of variance attributable to Level 1 (child) and Level 2 (classroom) for each outcome. For all predictive models, centering decisions were made according to recommendations by Enders and Tofighi (2007).

In Model 1, mean CLASS scores were grand mean centered and entered as Level 2 predictors, to examine whether mean classroom interaction quality was associated with children's behavior problems.

$$\text{Level 1: } \textit{Spring Behavior Problems}_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Quality Mean}) + u_{0j}$$

In Model 2, the respective variability scores were grand mean centered, to examine the association between variability in classroom interaction quality and behavior problems in the spring, above and beyond mean classroom interaction quality:

$$\text{Level 1: } \textit{Spring Behavior Problems}_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Quality Mean}) + \gamma_{02} (\text{Quality Variability}) + u_{0j}$$

In Model 3, the respective fall behavior problem and child demographic variables (age, sex, race and ethnicity) were entered as contextual variables, both at Level 1 and Level 2 (to obtain the Level 2 scores, individual scores within each *cluster*, or classroom, were aggregated). Level 1 fall behavior problem was grand mean centered and the random effects were freely estimated. Child age was continuous and was grand mean centered; the other child demographic variables (sex, race and ethnicity) were dichotomous (female, black and Hispanic were coded as 1, respectively). The random effects for all of the Level 1 child demographic covariates were fixed. In this final model, the unique association between variability in classroom interaction the quality and behavior problems in the spring was examined, controlling for mean classroom interaction quality, fall problem behavior and child demographic characteristics:

$$\text{Level 1: } \textit{Spring Behavior Problems}_{ij} = \beta_{0j} + \beta_{1j} (\text{Age}) + \beta_{2j} (\text{Sex}) + \beta_{3j} (\text{Race})$$

$$+ \beta_{4j} (\text{Ethnicity}) + \beta_{5j} (\text{Fall Problem Behavior}) + r_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Quality Mean}) + \gamma_{02} (\text{Quality Variability}) + \gamma_{03} (\text{Age})$$

$$+ \gamma_{04} (\text{Sex}) + \gamma_{05} (\text{Race}) + \gamma_{06} (\text{Ethnicity}) + \gamma_{07} (\text{Fall Problem Behavior})$$

$$+ u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50} + u_{5j}$$

Chapter 3: Results

Descriptive Statistics

Table 1 presents descriptive statistics for child- and classroom-level variables. Data were examined for homoscedasticity, skewness and kurtosis. No assumptions were found to be violated.

Psychometric Properties of Mean Classroom Interaction Quality Scores

RQ 1: Does mean classroom interaction quality vary as a function of primary activity setting? Table 2 provides the percent of observations for each activity setting and the mean classroom interaction quality for the three CLASS domains across the activity settings. The majority of observations were conducted during Small Group (33%), Whole Group (28%), and Free Play (27%); very few observations were conducted during Meals (9%) and Routines/Transitions (3%).

For the domain of Emotional Support, means varied significantly across activity settings, $F(4, 508) = 6.51, p < .001$. Pairwise comparisons indicated that Emotional Support was significantly higher during Free Play ($M = 5.73$) than all other activity settings (means ranged from 5.33 to 5.18). The mean for Classroom Organization also significantly varied across activity settings, $F(4, 508) = 7.86, p < .001$. During Free Play ($M = 5.24$), Classroom Organization was significantly higher during Free Play than the other activity settings (with means ranging from 5.00 to 4.41), except Small Group ($M = 5.07$). In addition, Classroom Organization means during both Small Group ($M = 5.07$) and Whole Group ($M = 5.00$) were significantly higher than during Meals ($M = 4.25$) and Routines/Transitions ($M = 4.41$). Instructional Support did not significantly differ across activity setting, overall, $F(4, 508) = 1.29, p = .27$. In sum, ANOVA findings indicate that

mean classroom interaction quality varied as a function of the primary activity setting for Emotional Support and Classroom Organization, but not for Instructional Support.

RQ 2: Are mean classroom interaction quality scores reliable across the various activity settings? Before examining the reliability of scores across activity settings, the rank ordering across cycle, or time was tested. The coefficients indicated adequate reliability (.83, .63, and .73 for Emotional Support, Classroom Organization, and Instructional Support respectively).

Table 3 presents the activity setting reliability coefficients for one through four cycles of observations. For all three CLASS domains, the G theory analyses resulted in adequate reliable estimates for four observation cycles (.91, .87, & .85) for Emotional Support, Classroom Organization and Instructional Support, respectively. Findings suggest that when four observation cycles are conducted, as delineated in the CLASS manual, the mean classroom interaction quality scores obtained are reliable across the various activity settings.

Unique Associations between Variability in Classroom Interaction Quality and Behavior Problems

RQ 3: What is the unique association between variability in classroom interaction quality and behavior problems, above and beyond mean classroom interaction quality? Table 4 presents bivariate correlations between child- and classroom-level variables. Variability in Emotional Support was positively associated with both Disruption and Disconnection. Variability in Classroom Organization was not significantly associated with Disruption or Disconnection. Variability in Instructional Support was negatively associated with both Disruption and Disconnection. Overall,

findings indicate that higher variability in Emotional Support was associated with higher behavior problems at the end of the school year; whereas higher variability in Instructional Support was associated with lower behavior problems in the spring.

Multilevel modeling results for Disruption. Table 5 presents the predictive model results for Disruption. The unconditional model suggested that 85.1% of the variance in Disruption in the spring was attributable to differences between children (Level 1) and 14.8% was attributable to differences between classrooms (Level 2).

In Model 1, mean classroom interaction quality scores were entered at Level 2, separately for each of the CLASS domains. Mean Emotional Support and Classroom Organization were negatively associated with Disruption and accounted for 8.4% and 17.3% of the Level 2 variance, respectively. Mean Instructional Support was not significantly associated with Disruption. Results indicate that higher mean Emotional Support and Classroom Organization predicted fewer disruptive behavior problems, on average, at the end of the school year.

In Model 2, the respective variability scores were entered into the model at Level 2. Variability in Emotional Support was positively associated with Disruption and accounted for an additional 10% of the Level 2 variance, after controlling for mean Emotional Support. Higher variability in Emotional Support predicted more disruptive problem behaviors in the spring, on average, above and beyond mean levels of Emotional Support. In this model, the association between mean Emotional Support and Disruption was no longer significant. Variability in Instructional Support was negatively associated with Disruption and accounted for 15% of the Level 2 variance, after controlling for mean Instructional Support. Higher variability in Instructional Support predicted fewer

disruptive problem behaviors in the spring, on average, above and beyond mean levels of Instructional Support. Variability in Classroom Organization was not significantly associated with Disruption. In this model, the negative association between mean Classroom Organization and Disruption remained significant. Higher mean Classroom Organization predicted fewer disruptive behavior problems, on average, at the end of the school year, after controlling for Classroom Organization variability.

In Model 3, child demographic covariates (age, sex, race and ethnicity) and the respective fall Disruption scores were entered as contextual variables, both at Level 1 and 2. For all three models, Emotional Support, Classroom Organization and Instructional Support, only Level 1 child sex and fall Disruption were significantly associated with Disruption in the spring. Girls displayed significantly lower Disruption in the spring than did boys and children with higher fall Disruption, exhibited higher Disruption in the spring.

Multilevel modeling results for Disconnection. Table 6 presents the predictive model results for Disconnection. The unconditional model indicated that 59.9% of the variance in Disconnection in the spring was attributable to differences between children (Level 1) and 40.1% was attributable to differences between classrooms (Level 2).

In Model 1, mean classroom interaction quality scores were entered at Level 2, separately for each of the CLASS domains. Mean Emotional Support and Classroom Organization were negatively associated with Disconnection and accounted for 6.6% and 9.8% of the Level 2 variance, respectively. Mean Instructional Support was not

significantly associated with Disconnection. Results indicate that higher mean Emotional Support and Classroom Organization predicted fewer disconnected behavior problems, on average, at the end of the school year.

In Model 2, the respective variability scores were entered into the model at Level 2. Variability in Emotional Support was positively associated with Disconnection and accounted for an additional 4% of the Level 2 variance, after controlling for mean Emotional Support. Higher variability in Emotional Support predicted more disconnected problem behaviors in the spring, on average, above and beyond mean levels of Emotional Support. In this model, the association between mean Emotional Support and Disruption was no longer significant. Variability in Instructional Support was negatively associated with Disconnection and accounted for 5.6% of the Level 2 variance, after controlling for mean Instructional Support. Higher variability in Instructional Support predicted fewer disconnected problem behaviors in the spring, on average, above and beyond mean levels of Instructional Support. Variability in Classroom Organization was not significantly associated with Disconnection. In this model, the negative association between mean Classroom Organization and Disconnection remained significant. Higher mean Classroom Organization predicted fewer disconnected behavior problems, on average, at the end of the school year, after controlling for Classroom Organization variability.

In Model 3, child demographic covariates (age, sex, race and ethnicity) and the respective fall Disconnection scores were entered as contextual variables, both at Level 1 and 2. For all three models, Level 1 fall Disconnection was positively associated with Disconnection in the spring. In addition, for the Emotional Support and Classroom Organization models, Level 2 race was positively associated with Disconnection;

children in classrooms with higher proportions of black children, displayed higher Disconnection, on average, in the spring. No other associations were significant.

Chapter 4: Discussion

The present study examined the role of variability in classroom interaction quality two ways. First, the psychometric properties of the typically used classroom quality mean scores were examined to test whether the levels (mean differences) and rank ordering (reliability) of the scores varied across activity settings. Second, variability scores were calculated for each of the three classroom quality domains to examine the unique associations between variability in classroom interaction quality and behavior problems, above and beyond mean classroom interaction quality. Findings provide additional evidence for the reliability of the CLASS and contribute to our understanding of the important role of variability in classroom interaction quality for children with behavior problems in Head Start.

Level of Classroom Interaction Quality Across Activity Settings

The hypotheses for the first research question were partially supported. Mean classroom interaction quality varied as a function of activity setting for Emotional Support and Classroom Organization, but not for Instructional Support. These findings are supported by prior research and bioecological theory, which suggest that varying classroom contexts place unique demands and challenges on teachers and children (Bronfenbrenner & Morris, 2006; Howes & Smith, 1995; Kontos & Wilcox-Herzog, 1997; Vitiello et al., 2012). Findings from the present study suggest that with these varying demands, the classroom interaction quality that children experience also significantly varies.

Findings identified certain activity settings in which the mean Emotional Support and Classroom Organization was significantly higher than other activity settings. For

classrooms to score high on Emotional Support, teachers must be warm, sensitive and responsive to the emotional needs of the children in their class (Pianta et al., 2008). Teachers in classrooms with high Classroom Organization structure and manage classroom activities, routines, and materials to support children's self-regulation, attention, and engagement in learning throughout the day (Pianta et al., 2008). For Emotional Support, the mean was significantly higher in Free Play than all other activity settings. Similarly, the Classroom Organization mean was higher during Free Play than the other activity settings, except Small Group. Classroom Organization during both Small Group and Whole Group was significantly higher than during Meals and Routines/Transition.

To understand these findings, it is important to consider the definition of developmentally appropriate practices in preschool and the extent to which each activity setting provides opportunities for such practices. Developmentally appropriate practice (DAP) is defined as child-focused, language-rich, multisensory experiences where children learn through their purpose-driven exploration and interaction with peers and teachers (Bredekamp & Copple, 1997). DAP also defines the role of the teachers as facilitative, responsive, supportive, and informative (Bredekamp & Copple, 1997). During free play, children are typically able to choose where and with what materials they would like to play. In a typical classroom during free play, children can be observed in various interest areas throughout the classroom. For example, there might be some children in the block area, with a few collaboratively building a structure with large blocks and others playing with dinosaurs; a few children in the dramatic play area, with a some acting out a story that was read in class earlier, using props and wearing costume,

while another child pretends to bathe a baby doll; a couple of children might be at the sand table, filling and dumping containers of varying sizes; several children can be observed completing puzzles, lacing beads or stacking pegs independently at a table; a couple of children painting at the easels in the art area; a few at the writing area, with one tracing letters on a small chalkboard and the others writing their names on blank pages; and a couple browsing books in the library area. Meanwhile, DAP states that the role of the teacher during free play is to observe and engage with children in the different interest areas to ideally extend the cognitive complexity of their play, increase understanding and prevent conflicts, by asking questions, acknowledging effort and cooperation, engaging in shared activities and conversations (Bredekamp & Copple, 1997). During other settings such as small and whole group, teachers lead the lessons and experiences, while children are typically expected to sit either at a table (during small group) or on the carpet (during whole group) and attend to and participate in the planned lesson. Routines, transitions and mealtimes are also typically more teacher-directed.

It is possible that because free play is generally a more flexible setting, teachers may be afforded more opportunities to engage in emotionally supportive interactions with children; whereas in more structured settings, teachers may focus more on fostering academic skills (during small or whole group) or on the health and safety of children (during routines, transitions and mealtimes). The flexibility of free play may also provide behavioral expectations that are more developmentally appropriate for preschool children, in that they are not expected to sit still, be quiet and attend for extended periods of time as they are in small and large group settings. Additionally, since children are able to choose what they want to play with during free play, they may select materials and

topics that are more relevant and interesting to them, leading to increased engagement and better behavior. Although teachers may consider children's interests, ask questions to facilitate engagement and provide hands on opportunities during small and large group times, it could be challenging to cater to and maintain the interests of all children.

Therefore, it is necessary to consider that the quality of interactions may be influenced not only by the nature of the activity setting, but by children's behaviors within each setting. In fact, research suggests that children's behavior varies across contexts (Kontos & Keyes, 1999; Vitiello et al., 2012; Qi, Kaiser, Milan, 2006). For example, children tend to be more disruptive during teacher- versus child-centered contexts (Qi et al., 2006). It is also well-documented that transition times are particularly challenging for children and teachers (Burden, 2003; Burts, Hart, Charlesworth & Kirk; Stainback & Stainback, 1996; Vitiello et al., 2012). If children are more disruptive in a teacher-centered setting, such as small and large group or during a transition, this might decrease the quality of the emotional connection between teachers and children, as measured by the Emotional Support domain.

Similarly, scores in the Classroom Organization domain might be particularly influenced by children's behaviors. In fact, one dimension, Behavior Management, directly assesses children's behavior and compliance (Pianta et al., 2008). Therefore, if children's behavior problems are more prevalent during some activity settings, this might contribute to lower mean scores.

Contrary to the hypothesis, mean classroom interaction quality did not vary as a function of activity settings for Instructional Support. Classrooms with high Instructional Support are characterized by interactions where teachers stimulate higher-order thinking

and problem-solving, provide appropriate feedback, engage in in-depth discussions, make learning meaningful to children and model the use of advanced language (Pianta et al., 2008). These are strategies that are difficult for teachers to implement, as suggested by overall low mean Instructional Support scores in national (publically funded pre-k and Head Start) and local samples (e.g., National Center for Development and Learning, 2007). In the present sample, the mean Instructional Support ranged from 2.08 to 2.66, indicating a very small (.57) difference between the highest (small group) and lowest (meals) rated setting. These overall low scores might have made it difficult to identify variability in mean levels of the quality of the Instructional Support across activity settings.

In sum, previous research has suggested that children's behavior differs across classroom contexts (Kontos & Keyes, 1999; Vitiello et al., 2012; Qi et al., 2006). The present study extends this literature by providing a more in depth examination of the classroom interaction quality that children experience within the various activity settings throughout a typical preschool morning. Results indicate that the emotional connection between teachers and children and the organization of preschool classrooms (as assessed by the Emotional Support and Classroom Organization domains of the CLASS, respectively) vary significantly as a function of activity settings. Further, Free Play was identified as the activity setting in which the mean level of emotional and organizational quality was highest.

Reliability of Mean Classroom Interaction Quality Across Activity Settings

Contrary to the hypothesis, mean classroom interaction quality scores were adequately reliable across activity settings for all three CLASS domains (Emotional

Support, Classroom Organization, and Instructional Support). This finding adds to the literature on the reliability of the CLASS scores, which have been found to be reliable in national, Head Start and dual language learner samples (Downer, et al., 2012; La Paro et al., 2004; Pianta et al., 2005). The present findings suggest that scores are also reliable across cycles and activity settings.

Before reliability could be examined across activity settings, it was necessary to test whether the rank ordering was consistent across cycle, or time. Results indicated that reliable coefficients were adequate for all three CLASS domains. This finding comports with previous research, which suggests that classroom interaction quality scores were relatively stable over time, across a typical morning (Curby, Grimm, & Pianta, 2010).

Despite the significant mean differences that were found across activity settings, results from the G-theory analyses suggest that scores were reliable across activity settings. In other words, classrooms rated as having higher classroom interaction quality in one activity setting are likely to have higher classroom interaction quality in the other activity settings, compared to other classrooms. It can be concluded that from a psychometric perspective, mean scores are sound and provide a reliable representation of the classroom interaction quality.

Association Between Classroom Interaction Quality and Behavior Problems

In partial support of the hypothesis, variability in Emotional Support and Instructional Support was associated with Disruption and Disconnection, above and beyond the respective mean classroom interaction quality. For Classroom Organization, the mean, but not variability, predicted problem behavior. However, in all of the predictive models, the associations were no longer significant once child-level

demographic variables and fall behavior problems were accounted for in the final multilevel models.

Mean classroom interaction quality. Emotional Support and Classroom Organization means were negatively associated with Disruption and Disconnection in the spring, whereas Instructional Support was not significantly associated with either of the two outcomes. These findings comport with the literature supporting within domain specificity of associations, such that there is a stronger prediction to child outcomes when the components of quality that are measured more closely align with the developmental domain (Downer, et al., 2010). For example, stronger associations have been found when instructional support is used to predict academic outcomes, than when emotional support is used; in addition, stronger associations have been found when emotional support is used to predict social-emotional outcomes, than when instructional support is used (Downer et al., 2010). Classroom Organization has been found to relate to both academic and social-emotional outcomes, which makes sense given the importance of self-regulation, attention, and engagement, in supporting both academic learning as well as social-emotional development (Downer et al., 2010).

Variability in classroom interaction quality. Variability in Emotional Support was positively associated, whereas variability in Instructional Support was negatively associated with Disruption and Disconnection, on average, in the spring, after accounting for mean classroom interaction quality. These findings indicate that in classrooms with high variability in Emotional Support, children, on average, exhibited more behavior problems at the end of the school year. Conversely, in classrooms with high variability in Instructional Support, children displayed fewer disruptive and disconnected behavior

problems, on average, in the spring. An interpretation of this set of findings is that variability in classroom interaction quality may not be universally “negative” when experienced by children with behavior problems. Rather the association between variability in classroom interaction quality and behavior problems may depend on the nature of the observed classroom interactions.

The positive association between variability in Emotional Support and behavior problems can be understood within the framework of attachment theory. The Emotional Support domain focuses on the emotional connection between teachers and children. Attachment theory posits that consistency in such emotional connections between children and their caregivers leads to secure attachments, which fosters positive social and cognitive development and influences the quality of future relationships (Ainsworth, 1969; Bowlby, 1969). In fact, prior research suggests that secure attachment is positively associated with social competence and negatively associated with ratings of behavior problems (Cutrona et al., 1994). Findings from the present study indicate that children who do not experience this consistency in the emotional connection with their teachers, but rather experience variability in this domain, exhibit higher behavior problems. This finding also comports with the early intervention literature, which suggests that predictability within the classroom (e.g., schedule, daily routine, rules) is critical for reducing children’s behavior problems (Fox et al., 2003; Powell et al., 2006). Finally, it replicates and extends the findings in the recent study by Curby and colleagues (2013), which found a positive association between consistency in preschool Emotional Support and social competence in kindergarten.

Variability in Instructional Support, on the other hand, was associated with fewer behavior problems. This finding is less intuitive, but perhaps may be understood by considering Vygotsky's zone of proximal development (Vygotsky, 1978). Vygotsky explains that for teachers to be most effective, they must provide instruction that is just above what children can accomplish on their own, but just below the point where it is too difficult, which may lead to frustration. This is challenging because children develop at different rates. In Head Start classrooms, in which there are typically 20 children of mixed ages (3- 5 years), one can imagine that children may vary greatly in their individual abilities.

Research suggests that when instruction is provided at children's ability level they are less likely to become frustrated and more likely to be engaged and attentive (Bordrova & Leong, 2005). Further, when children exhibit behaviors such as attention and engagement, they are less likely to display behavior problems (Dominguez & Greenfield, 2009; McWayne & Cheung, 2009). It is possible that teachers in classrooms with higher Instructional Support variability scores might be individualizing the level of instruction to meet the varying needs of the children in the class, thereby reducing children's likelihood for frustration and behavior problems.

Comporting with prior research findings supporting within domain specificity of the CLASS domains and child outcomes (Downer et al., 2010), mean Instructional Support was not significantly associated with the social-emotional outcomes. However, higher variability in Instructional support was associated with lower behavior problems in the spring. In light of previous research, the unique associations between variability in

Instructional Support and behavior problems were somewhat surprising. This finding suggests that variability in this domain may play a different role for children with behavior problems, than the mean.

Variability in Classroom Organization was not associated with children's behavior problems. Instead, mean Classroom Organization was predictive of both Disruption and Disconnection in the spring. Upon closer inspection of the items, it is evident that this domain inherently focuses to some degree on consistency. For example, observers are instructed to note whether teachers: provide concise, explicit and consistent behavioral expectations; enforce rules consistently; and effectively and consistently redirect misbehavior. Therefore, the mean score may already capture variability. For a classroom to score high on the mean Classroom Organization, the indicators for that domain must be consistently observed.

Controlling for child-level covariates. It is important to note that for all models, none of the associations between classroom interaction quality variability and/or or means remained significant once child-level demographic covariates and fall behavior problems accounted for. In fact, the respective fall behavior problem scores were the strongest predictors of children's behavior problems in the spring for all models. This is not surprising, because previous research has reported the difficulty in detecting main effects for classroom-level variables, when child-level predictors are included in the model (Dominguez et al., 2011).

For Disruption, child-level sex was significantly predictive of Disruption in the spring, with girls exhibiting fewer disruptive behavior problems than boys. This finding comports with prior research in Head Start, which suggests that preschool boys exhibit

more problem behavior in the context of peer play than girls (Coolahan, Fantuzzo, Mendez, & McDermott, 2000; Lutz, Fantuzzo, & McDermott, 2002). For Disconnection, classroom-level race was significantly predictive of spring behavior problems in the Emotional Support and Classroom Organization models. This indicates that children in classrooms with higher proportions of black children exhibited more disconnected behavior problems, on average.

In sum, results from the multilevel analyses suggest that variability in classroom interaction quality differentially relates to children's social-emotional outcomes, above and beyond mean levels. In fact, when variability scores were included in the models for Emotional Support, they explained additional variance in children's behavior problems and mean level associations were no longer significant. Further, mean Instructional Support was not associated with either of the two outcomes, but variability in Instructional Support, controlling for mean was significantly associated with both Disruption and Disconnection. It can be concluded that for at least two of the CLASS interaction quality domains (Emotional Support and Instructional Support), variability may be an important factor to be considered with respect to children's social-emotional outcomes.

Limitations and Future Directions

Despite the many contributions of the present study, there are several limitations that must be discussed. First, variability in classroom interaction quality was analyzed and discussed as a unique variable here. However, it is not completely independent from its respective mean levels. In fact, it was not appropriate to examine the interaction between variability and means because the correlations for the three domains were high.

It may be the case for all three domains, that the means somewhat capture variability. For Emotional Support, for example, the mean and variability were negatively associated. Visual inspection of the data plot for Emotional Support indicated that there were no instances of high variability and high means; similarly there were no instances of low variability and low means.

Second, there was a strong theoretical and empirical rationale for examining activity settings as a source of variability in classroom interaction quality means. However, other characteristics of classrooms (e.g., teacher-child ratio), teachers (e.g., perceptions of DAP; qualifications; stress) and children (e.g., behavior problems; disability classifications) that might contribute to variability in the quality of classroom interactions should be examined in future research. There may also be other characteristics of the CLASS measure that might contribute to variability between scores. For example, typically a primary activity setting as designated for each cycle, as was done in the present study; however, more than one setting could be captured within a single cycle. In addition, there might differences between raters that can contribute to variability among scores. Future studies should examine empirically whether these factors have an influence on classroom interaction quality means.

Third, directionality is being assumed here, in that classroom level variables were used to predict child outcomes. Despite research suggesting that teachers tend to have a greater influence on children's behavior than children's behavior has on teachers (Curby et al., 2014) very few cross-lagged analyses of these associations have been conducted. More research is needed to examine the directionality and magnitude of the potentially transactional relationship between teacher and child behaviors.

Finally, the present study was limited to social-emotional outcomes, namely, behavior problems within the context of peer play, in a sample of Head Start children. It is important to replicate these analyses with a broader set of child outcomes (e.g., academic domains) and in other samples. More rigorous empirical research is needed to further understand the role of variability in preschool classrooms and on children's developmental outcomes.

Implications for Policy and Practice

The present study has several implications for policy and practice. Broadly, it adds to the growing body of literature on the reliability and validity of the CLASS (Downer, et al., 2012; La Paro et al., 2004; Pianta et al., 2005). However, findings underscore the importance of looking beyond mean levels of classroom interaction quality, because such global ratings may miss important variations in quality that children experience across a typical morning in the preschool classroom.

From the present findings it can be concluded that the global ratings of classroom interaction quality (the overall mean level averaged across the four observation cycles, as recommended by the CLASS developers; Pianta, et al. 2008), which may include various activity settings, are sound psychometrically and provide a reliable representation of classroom interaction quality. Such scores are helpful in providing ratings of the average quality experienced in a classroom across a typical morning. However, findings from the present study also suggest that classroom interaction quality means vary as a function of activity setting.

Despite scores being reliable and stable in terms of rank, this does not guarantee that scores are equitable across classrooms if activity settings vary across classrooms.

The overall mean for a classroom can be affected by the activity settings that are included in the observations. Since the daily routine and schedule of activity settings vary from classroom to classroom, some teachers may be observed during activity settings that typically result in low scores (e.g., routines/transitions) while others may be observed during activity settings that typically result in high scores (e.g., free play).

This raises the issue of consequential validity (Messick, 1998). Programs, as well as funding and monitoring agencies are using these overall mean scores to make high-stakes decisions about teachers and programs. However, this is beyond the scope of the intended purpose of the measure, which calls into question the validity of inferences and conclusions made on the basis of such scores (Messick, 1998).

Within programs, there may be value for practitioners in conducting focused observations during specific activity settings to better inform program improvement and teacher professional development. For example, for early childhood coaches who observe teachers' classroom practice, using only global ratings (overall means) may make it difficult to pinpoint the source of a low score, especially because multiple activity settings may be captured within one observation cycle. Having descriptive data about the quality of interactions during specific activity settings may be more useful within early childhood educational programs to inform teacher professional development and monitor progress. In fact, the CLASS Implementation Guide (CLASS Implementation Guide, 2014) specifically states that scores should not be shared with teachers; instead objective examples of observations should be shared about one or two dimensions at a time.

However, it should be noted, that findings from the present study should not be taken to suggest that only focused observations or variability scores be used instead of the

overall CLASS mean scores for research or to inform practice. The overall global CLASS scores are still the most reliable scores when they are conducted over the four observation cycles. With focused observations, it is not guaranteed that each cycle will extend for 20 minutes (since many preschool activities are completed in a shorter period of time) and this may lower the reliability of the scores. In addition, the study findings provide evidence that variability scores are important to consider in understanding children's experiences in the Head Start classroom; however, much more research is needed to examine the extent to which this variability score could be used independently for research purposes and to inform practice. Without understanding overall mean levels, the nature and influence of variability on children's outcomes may be difficult to interpret. Until more research is conducted, the suggestion is that both should be used together to inform program improvement and teacher professional development, because both each provide valuable information about the classroom interaction quality that children experience.

In conclusion, findings broadly suggest that when using measures in early childhood programs, careful attention to reliability and validity should be paid. More specifically, results indicate that the classroom interaction quality that children experience varies across a typical morning and across activity settings. To illustrate this point, consider two children in classrooms with four cycles of observations (as are typically conducted): one in a classroom that is rated 4,4,4,4 -- consistent mid-level ratings across all four cycles -- and the other in classroom that is rated as 1,7,1,7 (although such extreme fluctuations are unlikely, it helps to illustrate the point). These classrooms both have the same global rating (overall mean) of 4, but the experiences of

the two children are likely very different. Findings from the present study suggest that understanding these patterns of variability are important because children's social-emotional outcomes may be influenced by this variability.

References

- Achenbach, T. M., & Edelbrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs of the Society for Research in Child Development*, 1-82.
- Achenbach, T. M., Edelbrock, C., & Howell, C. T. (1987). Empirically based assessment of the behavioral/emotional problems of 2-and 3-year-old children. *Journal of Abnormal Child Psychology*, 15(4), 629-650.
- Ainslie, R. C. (1990). Family and center contributions to the adjustment of infants in full-time day care. *New Directions for Child Development*, 49, 39 – 52.
- Belsky, J. (1997). Variation in susceptibility to rearing influences: An evolutionary argument. *Psychological Inquiry*, 8, 182–186.
- Belsky, J., Bakermans-Kranenburg, M.J., & Van Ijzendoorn, M.H. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science*, 16, 300–304.
- Bodrova, E. & Leong, D. J. (2005). High quality preschool programs: What would Vygotsky say? *Early Education and Development*, 16, 435–444.
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes.
- Bulotsky-Shearer, R., Fantuzzo, J. W., & McDermott, P. A. (2008). An investigation of classroom situational dimensions of emotional and behavioral adjustment and cognitive and social outcomes for Head Start children. *Developmental Psychology*, 44(1), 139-154.
- Bulotsky-Shearer, R. J., & Fantuzzo, J. W. (2011). Preschool behavior problems in classroom learning situations and literacy outcomes in kindergarten and first grade. *Early Childhood Research Quarterly*, 26(1), 61-73.
- Bulotsky-Shearer, R. J., Fernandez, V., Dominguez, X., & Rouse, H. L. (2011). Behavior problems in learning activities and social interactions in head start classrooms and early reading, mathematics, and approaches to learning. *School Psychology Review*, 40(1), 39-56.
- Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science*, 12(3), 140-153.
- Burts, D. C., Hart, C. H., Charlesworth, R. & Kirk, L. (1990). A comparison of frequencies of stress behaviors observed in kindergarten classrooms with developmentally appropriate versus developmentally inappropriate instructional practices. *Early Childhood Research Quarterly*, 5, 407–423.

- Campbell, S. B. & Ewing, L. J. (1990). Hard-to-manage preschoolers: Adjustment at age nine and predictors of continuing symptoms. *Journal of Child Psychology and Psychiatry*, 31,871-889.
- Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. *Development and Psychopathology*, 12(3), 467-488.
- CLASS Implementation Guide (2014). Teachstone. Retrieved from <http://teachstone.com/wp-content/uploads/2010/06/CLASSImplementationGuide.pdf>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Coolahan, K., Fantuzzo, J., Mendez, J., & McDermott, P. (2000). Preschool peer interactions and readiness to learn: Relationships between classroom peer play and learning behaviors and conduct. *Journal of Educational Psychology*, 92(3), 458-465.
- Curby, T. W., Brock, L. L., & Hamre, B. K. (2013). Teachers' Emotional Support Consistency Predicts Children's Achievement Gains and Social Skills. *Early Education & Development*, 24(3), 292-309.
- Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher-child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, 101(4), 912-925.
- Cronbach, L. J., Gieser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements*. New York: Wiley.
- Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? *Early Education and Development*, 17(1), 57-89.
- De Feyter, J. J. (2011). School readiness, early achievement, and the role of English language proficiency for children in low-income immigrant families (Doctoral dissertation, George Mason University). Retrieved from <http://hdl.handle.net/1920/6615>
- Dickinson, D.K., & Tabors, P.O. (Eds.). (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, MD: Brookes.
- Dobbs, J., Doctoroff, G. L., Fisher, P. H., & Arnold, D. H. (2006). The association between preschool children's socio-emotional functioning and their mathematical skills. *Journal of Applied Developmental Psychology*, 27(2), 97-108.

- Dominguez, X. & Greenfield, D. (2009). Learning behaviors mediating the effects of behavior problems on academic outcomes. *NHSA Dialog*, *12*, 1–17.
- Dominguez, X., Vitiello, V. E., Maier, M. F., & Greenfield, D. B. (2010). A Longitudinal Examination of Young Children's Learning Behavior: Child-Level and Classroom-Level Predictors of Change throughout the Preschool Year. *School Psychology Review*, *39*(1), 29-47.
- Dominguez, X., Vitiello, V. E., Fuccillo, J. M., Greenfield, D. B., & Bulotsky-Shearer, R. J. (2011). The role of context in preschool learning: A multilevel examination of the contribution of context-specific problem behaviors and classroom process quality to low-income children's approaches to learning. *Journal of School Psychology*, *49*(2), 175-195.
- Donohue, K. M., Perry, K. E., & Weinstein, R. S. (2003). Teachers' classroom practices and children's rejection by their peers. *Journal of Applied Developmental Psychology*, *24*(1), 91-118.
- Downer, J. T., López, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: Applicability of the Classroom Assessment Scoring System in diverse settings. *Early Childhood Research Quarterly*, *27*(1), 21-32.
- Downer, J., Sabol, T. J., & Hamre, B. (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*(5), 699-723.
- Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning?. *School Psychology Review*, *36*(3), 413-432.
- Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H., ... Duckworth, K. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428-1446.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121-138. doi:10.1037/1082-989X.12.2.121
- Fantuzzo, J. W., Bulotsky-Shearer, R., Fusco, R. A., & McWayne, C. (2005). An investigation of preschool classroom behavioral adjustment problems and social–emotional school readiness competencies. *Early Childhood Research Quarterly*, *20*(3), 259-275.
- Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review*, *36*, 44–62.

- Fantuzzo, J., Bulotsky, R., McDermott, P., Mosca, S., & Lutz, M. N. (2003). A multivariate analysis of emotional and behavioral adjustment and preschool educational outcomes. *School Psychology Review, 32*(2), 185–203.
- Fox, L., Dunlap, G., & Cushing, L. (2002). Early intervention, positive behavior support, and transition to school. *Journal of Emotional and Behavioral Disorders, 10*(3), 149-157.
- Gazelle, H. (2006). Class climate moderates peer relations and emotional adjustment in children with an early history of anxious solitude: A child X environment model. *Developmental Psychology, 42*(6), 1179-1192.
- Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. Foundation for Child Development.
- Goossens, F. A., & van IJzendoorn, M. H. (1990). Quality of infants' attachments to professional caregivers: Relation to infant – parent attachment and day-care characteristics. *Child Development, 61*, 832 – 837.
- 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.
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale*. Teachers College Press, Columbia University, 1234 Amsterdam Avenue, New York, NY 10027.
- Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21*(2), 143-150.
- Howes, C. (2011). Children's social development within the socialization context of child care and early childhood education. *The Wiley-Blackwell Handbook of Childhood Social Development, Second Edition*, 246-262.
- 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.
- Howes, C., & Smith, E. W. (1995). Relations among child care quality, teacher behavior, children's play activities, emotional security, and cognitive activity in child care. *Early Childhood Research Quarterly, 10*(4), 381-404.
- Huaqing, Q. C., & Kaiser, A. P. (2003). *Behavior problems of preschool children from low-income families: Review of the literature* PRO-ED, Inc.

- Johnson, T. C., Stoner, G., & Green, S. K. (1996). Demonstrating the experimenting society model with classwide behavior management interventions. *School Psychology Review, 25*(2), 199-214.
- Kern, L., & Clemens, N. H. (2007). Antecedent strategies to promote appropriate classroom behavior. *Psychology in the Schools, 44*(1), 65-75.
- Kontos, S., & Wilcox-Herzog, A. (1997). Influences on children's competence in early childhood classrooms. *Early Childhood Research Quarterly, 12*(3), 247-262.
- La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal, 104*(5), 409-426.
- Lamb, M. E., & Ahnert, L. (2006). Nonparental child care: Context, concepts, correlates and consequences. In W. Damon, R. M. Lerner, K. A. Renninger, & I. E. Sigel (Eds.), *Handbook of child psychology: Volume 4: Child psychology in practice (5th ed.)*. Hoboken, NJ: Wiley.
- Lonigan, C. J., Hooe, E. S., David, C. F., & Kistner, J. A. (1999). Positive and negative affectivity in children: Confirmatory factor analysis of a two-factor model and its relation to symptoms of anxiety and depression. *Journal of Consulting and Clinical Psychology, 67*(3), 374.
- Lutz, M. N., Fantuzzo, J., & McDermott, P. (2002). Multidimensional assessment of emotional and behavioral adjustment problems of low-income preschool children: Development and initial validation. *Early Childhood Research Quarterly, 17*(3), 338-355.
- Mashburn, A. J., & Pianta, R. C. (2006). Social relationships and school readiness. *Early Education and Development, 17*(1), 151-176.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., 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.
- McWayne, C., & Cheung, K. (2009). A picture of strength: Preschool competencies mediate the effects of early behavior problems on later academic and social adjustment for Head Start children. *Journal of Applied Developmental Psychology, 30*(3), 273-285.
- Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research, 45*(1-3), 35-44.

- National Association for the Education of Young Children (2009). Developmentally appropriate practice in early childhood programs serving children from birth through age 8: A position statement of the National Association for the Education of Young Children. Retrieved from <http://www.naeyc.org/files/naeyc/file/positions/PSDAP.pdf>
- National Education Goals Panel. (1997). *Special early childhood report*. Washington, DC: U.S. Government Printing Office.
- National Institute for Early Education Research (U.S.), & Pew Charitable Trusts. (2011). *The state of preschool: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.
- Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1), 1-18.
- O'Connor, E. E., Dearing, E., & Collins, B. A. (2011). Teacher-child relationship and behavior problem trajectories in elementary school. *American Educational Research Journal*, 48(1), 120-162.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system*. Baltimore: Paul H. Brookes.
- 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*, 225-238.
- Pluess, M., & Belsky, J. (2009). Differential susceptibility to rearing experience: The case of childcare. *Journal of Child Psychology and Psychiatry*, 50(4), 396-404.
- Powell, D., Dunlap, G., & Fox, L. (2006). Prevention and intervention for the challenging behaviors of toddlers and preschoolers. *Infants & Young Children*, 19(1), 25-35.
- Raver, C., Knitzer, J. (2002). Ready to enter: What research tells policymakers about strategies to promote social and emotional school readiness among three- and four-year-old children. *Promoting the emotional well-being of children and families policy paper* National Center for Children in Poverty, Columbia University.
- Raver, C. C., & Zigler, E. F. (1997). Social competence: An untapped dimension in evaluating Head Start's success. *Early Childhood Research Quarterly*, 12(4), 363-385.
- Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2000). *Hierarchical Linear and Nonlinear Modeling-HLM5*. Lincolnwood, IL: SSI Scientific Software.
- Richman, N., Stevenson, J., & Graham, P. J. (1982). *Preschool to school: a behavioral study*. London: Academic Press.

- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., 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*(4), 958-972.
- Rimm-Kaufman, S. E., Early, D. M., Cox, M. J., Saluja, G., Pianta, R. C., Bradley, R. H., & Payne, C. (2002). Early behavioral attributes and teachers' sensitivity as predictors of competent behavior in the kindergarten classroom. *Applied Developmental Psychology, 23*, 451-470.
- Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal, 105*(4), 377-394.
- Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology, 21*(5), 491-511.
- Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly, 15*(2), 147-166.
- Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., & Raudenbush, S. (2011). Optimal Design Plus Version 3.0.
- Thompson, R. A., & Raikes, H. A. (2007). The social and emotional foundations of school readiness. In D. F. Perry, R. K. Kaufmann, & J. Knitzer (Eds.), *Social and emotional health in early childhood: Building bridges between services and systems* (pp. 13–35). Baltimore: Brookes.
- Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly, 27*(2), 210-220.
- Vygotsky, L. S. (1978). Interaction between learning and development (M. Lopez-Morillas, Trans.). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes* (pp. 79–91). Cambridge, MA: Harvard University Press.
- Webster-Stratton, C., & Hammond, M. (1998). *Marital conflict management skills, parenting style and early-onset conduct problems: Processes and pathways*. Unpublished manuscript. Seattle, WA: University of Washington.
- Webster-Stratton, C., Reid, J., & Hammond, M. (2001). Social skills and problem-solving training for children with early-onset conduct problems: who benefits?. *Journal of Child Psychology and Psychiatry, 42*(7), 943-952.

Wilson, H. K., Pianta, R. C., & Stuhlman, M. (2007). Typical classroom experiences in first grade: The role of classroom climate and functional risk in the development of social competencies. *The Elementary School Journal*, *108*(2), 81-96.

Table 1

Descriptive Statistics

	<i>n</i>	<i>M</i>	<i>SD</i>
PIPPS (fall)			
Disruption with Peers	53	50.07	7.15
Disconnected with Peers	53	44.68	10.32
PIPPS (spring)			
Disruptive with Peers	50	50.16	7.20
Disconnected with Peers	51	43.93	10.09
CLASS Mean (winter)			
Mean Emotional Support	54	5.27	0.97
Mean Classroom Organization	53	4.59	0.98
Mean Instructional Support	54	2.48	1.04
CLASS Variability (winter)			
Variability in Emotional Support	53	1.03	0.38
Variability in Classroom Organization	52	0.84	0.26
Variability in Instructional Support	53	2.48	0.35

Note. Scores for the PIPPS represent standardized *T* scores ($M = 50$, $SD = 10$). Scores for the CLASS Mean represent average scores across cycles for each domain (Range 1-7).

Table 2

Activity Setting Means

	Free Play <i>M (SE)</i>	Whole Group <i>M (SE)</i>	Small Group <i>M (SE)</i>	Meals <i>M (SE)</i>	Routines/Transitions <i>M (SE)</i>	Significant Differences ^a
Emotional Support	5.73 (.08)	5.33 (.07)	5.27 (.08)	5.19 (.22)	5.18 (.14)	FP > WG, SG, M, RT
Classroom Organization	5.25 (.09)	5.00 (.08)	5.07 (.09)	4.25 (.26)	4.41 (.15)	FP > WG, M, RT WG, SG > M, RT
Instructional Support	2.57 (.09)	2.50 (.08)	2.66 (.09)	2.08 (.26)	2.58 (.15)	--

Percent of observations

27%

28%

33%

9%

3%

Note. Activity settings are denoted as follows: FP (Free Play); WG (Whole Group); SG (Small Group); M (Meals); and R/T (Routines/Transitions). Percent of observations provides the percent of observations which captured each of the activity settings.
^aStatistically different differences between Activity Setting means from ANOVA Pairwise Comparisons ($p < .05$).

Table 3

Activity Setting Generalizability Theory Results

	Classroom Variance (σ_c^2)	Child Variance (T)	Reliability Coefficients			
			1 Cycles $\sigma_a^2 / (\sigma_a^2 + \sigma_{ca,e}^2)$	2 Cycles $\sigma_a^2 / (\sigma_a^2 + \sigma_{ca,e}^2 / 2)$	3 Cycles $\sigma_a^2 / (\sigma_a^2 + \sigma_{ca,e}^2 / 3)$	4 Cycles $\sigma_a^2 / (\sigma_a^2 + \sigma_{ca,e}^2 / 4)$
Emotional Support	0.55	0.21	0.72	0.84	0.89	0.91
Classroom Organization	0.58	0.33	0.64	0.78	0.84	0.87
Instructional Support	0.60	0.42	0.59	0.74	0.81	0.85

Note. The variance in classroom interaction quality (CLASS) means was partitioned into variance attributed to individual difference between classrooms (σ_c^2), variance attributed to activity setting conditions (σ_a^2) and residual variance ($\sigma_{ca,e}^2$), which included the interaction between classroom (c) and activity setting (a), as well as random error (e) (Cronbach et al., 1972). Reliability coefficients are typically calculated for one instance of a variable: $\sigma_a^2 / (\sigma_a^2 + \sigma_{ca,e}^2)$. For the present study, four consecutive cycles of observations were conducted for each classroom (as delineated in the CLASS manual; Pianta et al., 2008). Therefore, reliability estimates were calculated to represent the reliability estimates for four cycles of observations: $\sigma_a^2 / (\sigma_a^2 + \sigma_{ca,e}^2 / 4)$.

Table 4

Bivariate Correlations between Behavior Problems and Classroom Interaction Quality

	1	2	3	4	5	6	7	8	9	10
1. Disruption (fall)	—	.63**	.72**	.50**	-.31*	-.34*	-.19	.45**	.13	-.31*
2. Disconnection (fall)		—	.36*	.67**	-.25	-.22	-.20	.30*	-.01	-.33*
3. Disruption (spring)			—	.61**	-.32*	-.36*	-.18	.43**	.12	-.39**
4. Disconnection (spring)				—	-.24	-.34*	-.16	.35*	.18	-.33*
5. Mean Emotional Support					—	.78**	.68**	-.74**	-.21	.58**
6. Mean Classroom Organization						—	.65**	-.70**	-.35*	.61**
7. Mean Instructional Support							—	-.55**	-.22	.70**
8. Variability in Emotional Support								—	.15	-.45**
9. Variability in Classroom Organization									—	-.14
10. Variability in Instructional Support										—

* $p < .05$, ** $p < .01$

Table 5

Multilevel Models for Variability in Classroom Interaction Quality predicting Disruptive Behavior Problems

Fixed Effects	Emotional Support			Classroom Organization			Instructional Support		
	Coefficient	df	t-ratio	Coefficient	df	t-ratio	Coefficient	df	t-ratio
Model 1									
Intercept (β_{0j})	50.06***	46	111.54	49.95***	46	115.73	49.91***	46	108.39
Mean (γ_{01})	-0.98*	46	-2.05	-1.25**	46	-2.75	-0.52	46	-1.77
Model 2									
Intercept (β_{0j})	50.05***	45	115.33	49.95***	45	114.50	50.01***	45	114.48
Mean (γ_{01})	-0.13	45	-0.21	-1.28*	45	-2.68	0.49	45	-2.54
Variability (γ_{02})	1.66*	45	2.12	-0.22	45	-0.24	-2.09*	45	0.84
Model 3									
Intercept (β_{0j})	58.12***	39	8.68	58.29***	39	8.80	57.66***	39	8.38
Mean (γ_{01})	-0.18	39	-0.55	-0.31	39	-1.19	-0.18	39	-0.60
Variability (γ_{02})	0.13	39	0.28	0.24	39	0.49	-0.14	39	-0.33
Age (γ_{03})	-0.09	39	-0.87	-0.07	39	-0.70	-0.07	39	-0.70
Sex (γ_{04})	-1.10	39	-0.40	-1.52	39	-0.56	-1.40	39	-0.51
Race (γ_{05})	1.40	39	1.10	1.74	39	1.34	1.27	39	0.99
Ethnicity (γ_{06})	1.10	39	0.85	1.34	39	1.08	0.87	39	0.72
Fall Disruption (γ_{07})	-0.08	39	-1.02	-0.10	39	-1.28	-0.08	39	-1.04
Age (β_{1j})	-0.01	760	-0.28	-0.01	760	-0.27	-0.01	760	-0.25
Sex (β_{2j})	-0.66*	760	2.11	-0.67*	760	-2.13	-0.67*	760	-2.15
Race (β_{3j})	0.61	760	0.83	0.61	760	0.82	0.62	760	0.84
Ethnicity (β_{4j})	-0.55	760	-0.75	-0.56	760	-0.78	-0.54	760	-0.74
Fall Disruption (β_{5j})	0.80***	760	16.95	0.79***	760	16.90	0.79***	760	16.97
Random Effects	Variance Component	df	χ^2	Variance Component	df	χ^2	Variance Component	df	χ^2
Intercept (τ_{00})	3.04***	39	138.10	3.95***	39	134.67	3.05	39	138.19
Fall Disruption (τ_{07})	0.05***	46	105.46	0.05***	46	105.47	0.5	46	105.49
Level-1 effects (σ^2)	17.68			17.67			17.67		

* $p < .05$. ** $p < .01$. *** $p < .001$.Note. Disruption scores represent standardized T scores ($M = 50$, $SD = 10$). Random effects are presented for the final model (Model 3).

Table 6

Multilevel Models for Variability in Classroom Interaction Quality predicting Disconnected Behavior Problems

Fixed Effects	Emotional Support			Classroom Organization			Instructional Support		
	Coefficient	df	t-ratio	Coefficient	df	t-ratio	Coefficient	df	t-ratio
Model 1									
Intercept (β_{0i})	43.93***	47	47.74	43.89***	47	48.50	43.83***	47	46.78
Mean (γ_{01})	-1.99*	47	-2.02	-2.32*	47	-2.42	-1.28	47	-1.41
Model 2									
Intercept (β_{0i})	44.07***	46	115.33	43.89***	46	114.50	43.97***	46	48.02
Mean (γ_{01})	-0.61	46	-0.48	-2.32*	46	-2.29	0.35	46	0.29
Variability (γ_{02})	2.67*	46	1.64	0.01	46	0.01	-3.41*	46	-1.97
Model 3									
Intercept (β_{0i})	44.86**	40	3.28	47.24	40	3.53	50.34***	40	3.70
Mean (γ_{01})	0.45	40	0.55	-0.50	40	-0.72	0.01	40	0.01
Variability (γ_{02})	1.38	40	1.13	1.03	40	0.81	0.10	40	0.08
Age (γ_{03})	-0.13	40	-0.50	-0.13	40	-0.52	-0.20	40	-0.79
Sex (γ_{04})	4.21	40	0.59	1.17	40	0.17	1.83	40	0.26
Race (γ_{05})	6.14*	40	2.14	6.25*	40	2.13	5.40	40	1.75
Ethnicity (γ_{06})	3.62	40	1.19	3.25	40	1.12	2.07	40	0.72
Fall Disruption (γ_{07})	-0.00	40	-0.03	-0.01	40	-0.13	0.01	40	0.09
Age (β_{1i})	-0.05	763	-1.33	-0.05	763	-1.34	-0.05	763	-1.34
Sex (β_{2i})	-0.63	763	-1.39	-0.63	763	-1.39	-0.63	763	-1.39
Race (β_{3i})	-0.40	763	-0.37	-0.40	763	-0.38	-0.39	763	-0.37
Ethnicity (β_{4i})	-2.00	763	-1.90	-2.01	763	-1.91	-2.00	763	-1.90
Fall Disruption (β_{5i})	0.60***	763	19.04	0.60***	763	19.04	0.60***	763	19.03
Random Effects	Variance Component	df	χ^2	Variance Component	df	χ^2	Variance Component	df	χ^2
Intercept (τ_{00})	3.45***	40	205.28	3.46***	40	199.28	3.50	40	203.35
Fall Disruption (τ_{07})	0.17***	46	86.16	0.17	46	86.20	0.17	46	86.27
Level-1 effects (σ^2)	5.98			5.98			5.98		

* $p < .05$. ** $p < .01$. *** $p < .001$.Note. Disconnection scores represent standardized T scores ($M = 50$, $SD = 10$). Random effects are presented for the final model (Model 3).