

2019-07-31

Examining the Role of Varying Levels of Classroom Quality for Toddlers in Early Head Start and Subsidized Child Care Programs: Understanding Threshold Effects

Krystal Bichay Awadalla
University of Miami, kbichay@miami.edu

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

Recommended Citation

Bichay Awadalla, Krystal, "Examining the Role of Varying Levels of Classroom Quality for Toddlers in Early Head Start and Subsidized Child Care Programs: Understanding Threshold Effects" (2019). *Open Access Dissertations*. 2351.
https://scholarlyrepository.miami.edu/oa_dissertations/2351

This Open access 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 ROLE OF VARYING LEVELS OF CLASSROOM QUALITY FOR
TODDLERS IN EARLY HEAD START AND SUBSIDIZED CHILD CARE
PROGRAMS: UNDERSTANDING THRESHOLD EFFECTS

By
Krystal Bichay Awadalla

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

August 2019

©2019
Krystal Bichay Awadalla
All Rights Reserved

UNIVERSITY OF MIAMI

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

EXAMINING THE ROLE OF VARYING LEVELS OF CLASSROOM QUALITY FOR
TODDLERS IN EARLY HEAD START AND SUBSIDIZED CHILD CARE
PROGRAMS: UNDERSTANDING THRESHOLD EFFECTS

Krystal Bichay Awadalla

Approved:

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

Daryl Greenfield, Ph.D.
Professor of Psychology

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

Elizabeth A. Simpson, Ph.D.
Assistant Professor of Psychology

Ann-Marie Faria, Ph.D.
Principal Researcher
American Institutes for Research

Guillermo Prado, Ph.D.
Dean of the Graduate School

AWADALLA, KRYSTAL BICHAY

(Ph.D., Psychology)

Examining the Role of Varying Levels of Classroom Quality for Toddlers in Early Head Start and Subsidized Child Care Programs: Understanding Threshold Effects.

(August 2019)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Rebecca Bulotsky Shearer.
No. of pages in text. (82)

The Toddler Classroom Assessment Scoring System (CLASS-T) is one of the most commonly used measures to assess the quality of teacher-child interactions in toddler classrooms. Cut-points on the CLASS-T scores are being used to determine Quality Rating and Improvement Systems (QRIS) ratings of programs as high or low quality, and therefore have policy implications for program funding and for parent consumers. Despite widespread use of the CLASS-T for these purposes, few studies have examined the construct validity of the CLASS-T for use in Early Head Start (EHS) and subsidized child care programs serving low-income, ethnically and linguistically diverse children and families. In addition, no studies, to date, have validated a priori cut-points that QRIS programs use in practice. Therefore, the purpose of the present study was (a) to examine the factor structure of the CLASS-T in a sample of 106 classrooms comprised of 688 ethnically and linguistically diverse toddlers attending EHS and subsidized child care programs; (b) to examine associations between the CLASS-T domain scores and child developmental skills, including language and social-emotional skills; and (c) to empirically examine the cut-points that were in use as part of the local Miami-Dade County QRIS,

Quality Counts, to determine whether the a priori cut-points used in practice aligned with thresholds of classroom quality that were differentially associated with child developmental skills. Findings supported the two-factor structure of the CLASS-T, which included the following two factors: Emotional and Behavioral Support, and Engaged Support for Learning. Both linear and nonlinear relationships between CLASS-T domain scores and social-emotional skills were found, and the CLASS-T a priori cut-points used by Quality Counts were aligned with thresholds in the statistical models examining social-emotional skills. The findings of the current study were shared with local partners and implications for policy and practice are discussed.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	v
 Chapter	
1 INTRODUCTION	1
Theoretical Framework.....	4
Classroom Quality.....	5
Classroom Quality in Toddler Classrooms.....	7
Associations between Process Quality and Child Developmental Skills	9
Thresholds of Classroom Process Quality	10
The Current Study.....	15
Research Objectives and Hypotheses	15
2 METHOD	18
Participants.....	18
Procedures.....	19
Measures.....	20
Data Analytic Plan.....	23
3 RESULTS	32
Descriptive Statistics.....	32
Objective 1: Construct Validity of the CLASS-T.....	32
Objective 2: Linear Relationship Between the CLASS-T and Child Developmental Skills	34
Objective 3: Thresholds in the Relationship Between the CLASS-T and Child Developmental Skills.....	36
4 DISCUSSION	42
Construct Validity of the CLASS-T	42
Linear Relationship Between the CLASS-T and Child Developmental Skills	45
Thresholds in the Relationship Between the CLASS-T and Child Developmental Skills	47
Limitations and Future Directions	54
Implications for Policy and Practice	57
REFERENCES.....	61
FIGURES.....	70
TABLES.....	77

LIST OF FIGURES

Figure 1. Factor Loadings from Final Two-Factor Model of CLASS-T with Overall Sample	70
Figure 2. Factor Loadings from CFA of Two-Factor Model of CLASS-T in Classrooms that Use Predominantly English.....	71
Figure 3. Factor Loadings from CFA of Two-Factor Model of CLASS-T in Classrooms that Use Predominantly Spanish.....	72
Figure 4. Nonlinear Relationship between EBS and BITSEA Social-emotional Problems	73
Figure 5. Nonlinear Relationship between ESL and BITSEA Social-emotional Problems	74
Figure 6. Nonlinear Relationship between EBS and BITSEA Social-emotional Competence.....	75
Figure 7. Nonlinear Relationship between ESL and BITSEA Social-emotional Competence.....	76

LIST OF TABLES

Table 1. Means and Standard Deviations for Measures.....	77
Table 2. Bivariate Correlations between Measures.....	78
Table 3. Linear Relationships between CLASS-T and BITSEA Domains	79
Table 4. Linear Relationships between CLASS-T and ASQ-3 Domains.....	80
Table 5. Testing for Quality Thresholds using A Priori Cut-points of CLASS-T...	81
Table 6. Testing for Quality Thresholds Empirically using B-spline Analyses: Nonlinear F Tests	82

CHAPTER 1 INTRODUCTION

At this time of rapid expansion of early care and education (ECE) programs to reach children from birth to age three, ensuring the provision of high quality experiences to *all children* is a national policy issue. The number of young children, between the ages of 15 and 36 months, enrolled in center-based ECE has increased, with 34% of children between the ages of 2 and 3 enrolled in either center- or home-based care (U.S. Census Bureau, 2013). Many of these children are from low-income backgrounds, with 7.8 million children under the age of three living in poverty (Jiang, Granja, & Koball, 2017a). Research suggests that the daily stressors that result from living in poverty are associated with an achievement gap in reading and mathematics at kindergarten entry (Hammer, Farkas, & Maczuga, 2010; Hindman, Skibbe, Miller, & Zimmerman, 2010; Magnuson & Duncan, 2006). Fortunately, research shows that high quality ECE experiences for toddlers from low-income families buffer the negative risks to achievement that are associated with living in poverty (Dearing, McCartney, & Taylor, 2009; McCartney, Dearing, Taylor, & Bub, 2007; Sabol & Pianta, 2012; Watamura, Phillips, Morrissey, McCartney, & Bub, 2011). In fact, research suggests that high quality ECE experiences in toddlerhood are associated with higher academic achievement and lower externalizing behavior problems in elementary school (Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010), making visible the importance of quality ECE experiences, particularly for toddlers from low-income families.

Unfortunately, research to date suggests that quality in toddler classrooms is low to mediocre, particularly in programs serving children from low-income families (Bandel, Aikens, Vogel, Boller, & Murphy, 2014; Ruzek, Burchinal, Farkas, & Duncan, 2014).

Children living in low-income families are less likely to have access to high quality care (Hatfield, Lower, Cassidy, & Faldowski, 2015). In addition, ethnically and linguistically diverse families face barriers to access quality care due to immigrant status, acculturation status and linguistic differences (Hernandez, Denton, & McCartney, 2011; Vesely, 2013).

Several recent national policy initiatives have responded to the need for quality early childhood programs to reach children and families living in poverty. These initiatives include the investment of 635 million dollars to increase access to quality ECE programs for infants and toddlers from low-income households through the Early Head Start (EHS)-Child Care partnerships and expansion grants (National Center on Early Head Start- Child Care Partnerships, 2016). Additionally, The Child Care and Development Block Grant Act of 2014 (P.L 113-186), reauthorized the Child Care and Development Fund (CCDF) program to expand access to quality care for children from low-income families eligible to receive child care subsidies (Office of Child Care, 2014). Policymakers have also increased the funding invested in quality improvement initiatives, including QRIS, to provide supports for quality improvement efforts, such as on-site technical assistance, professional development, and stipends (Tout, Zaslow, Halle, & Forry, 2009; Zaslow, Tout, & Martinez-Beck, 2010).

To guide quality improvement efforts, measures of *classroom process quality* are used. Process quality comprises dynamic aspects of the classroom, such as teacher-child interactions, that directly influence children's development and learning (Burchinal, Zaslow, & Tarullo, 2016). The Toddler Classroom Assessment Scoring System (CLASS-T; La Paro, Hamre, & Pianta, 2012) is one of the most commonly used tools to measure classroom process quality in toddler classrooms. Scores on the CLASS-T are used to

guide intervention supports and to monitor the effectiveness of quality improvement efforts in EHS and infant/toddler child care programs participating in QRIS (Hsu, 2017; Kisker et al., 2011). Despite the widespread use of the CLASS-T nationally, to date, few studies have examined the validity of the CLASS-T for use within EHS and subsidized child care programs serving ethnically and linguistically diverse toddlers.

In addition, within early childhood quality improvement initiatives such as QRIS, toddler programs receive ratings based on different components of quality, which include measures of *classroom process quality*, such as the CLASS-T (Schaack, Tarrant, Boller, & Tout, 2012). Cut-points based on scores from these measures of process quality are used to define “high” or “low” quality programs and may contribute to ratings used to determine funding (e.g., tiered reimbursement models for subsidized child care) or made public to parents to help them make informed child care decisions. These cut-points, however, are typically arbitrarily set by policy-makers or practitioners and not empirically validated by researchers (Pianta, Downer, & Hamre, 2016). Early childhood programs need empirical evidence that the cut-points used in practice to rate “high” and “low” quality programs actually align with thresholds that contribute to better developmental skills for young children. As ECE programs like EHS and subsidized child care programs increasingly enroll diverse populations of children and families, research is needed to identify empirically the point at which increases in classroom process quality contribute to children’s developmental skills within programs serving diverse, low-income families.

Despite widespread use of measures like the CLASS-T in early childhood programs to measure quality, several critical gaps currently exist. Few studies to date

have examined the reliability and validity of the CLASS-T in ethnically and linguistically diverse, low-income samples, and no studies, to date, have validated the underlying factor structure when the CLASS-T is used within diverse classrooms, such as those classrooms where teachers' predominantly use another language other than English. Additionally, there is inconsistent evidence for linear associations between the CLASS-T domain scores and child developmental skills in programs serving toddlers from low-income families. Furthermore, no studies, to date, have validated empirically the cut-points currently used by QRIS and other quality improvement initiatives that contribute to ratings of "high" and "low" quality in toddler programs serving low-income families.

To address these critical gaps, the objectives of the current study were to: (1) Extend the reliability and validity of the CLASS-T in a sample of classrooms comprised of ethnically and linguistically diverse toddlers from low-income families; (2) Examine the associations between process quality as measured by the CLASS-T and children's developmental skills, including social-emotional and language skills; and (3) Empirically validate the cut-points being used in practice within the local QRIS in an ethnically and linguistically diverse sample of toddlers from low-income families.

Theoretical Framework

Two broad theoretical frameworks underscore the importance of studying *classroom process quality*, and in particular teacher-child interactions, with toddlers. First, the bioecological model (Bronfenbrenner & Morris, 2006) highlights the importance of the proximal contexts that directly influence children's development, such as the classroom. Within the classroom, children's interactions with others in the

environment, particularly teachers, provide learning opportunities that support foundational developmental skills, including language and social-emotional skills.

Second, constructivist theory provides a framework to guide the study of process quality and learning within the toddler classroom context. This theory argues that children construct their own learning by interacting with others and materials in their natural environment (Bodrova, Leong, Dickinson, & Neuman, 2006; Piaget, 1951; Vygotsky, 1978). These interactions provide many opportunities to actively explore while learning (Copple & Bredekamp, 2009; Pianta, La Paro, & Hamre, 2008), which support the development of important language and social-emotional skills (Hamre et al., 2012; Mashburn et al., 2008). The framework posits that social interaction is necessary for learning and that teachers serve as toddlers' primary social partners. Within children's zone of proximal development, teachers scaffold children's experiences to help them learn and extend what they already know. Through interactions with teachers, toddlers are able to effectively explore their environments and learn as much as possible by having a secure base that enables them to feel comfortable and confident exploring the environment and learning.

Classroom Quality

Classroom quality is defined broadly as aspects of the environment and children's interactions that foster their physical, social, and cognitive development (Layzer & Goodson, 2006). Classroom quality is distinguished into two types: *structural* and *process quality*. *Structural quality* is characterized by aspects of the classroom that can be regulated, which include teacher education, teacher-child ratio, and group sizes (Cassidy et al., 2005; Hestenes, Cassidy, Hegde, & Lower, 2007; La Paro et al., 2012; Pianta et al.,

2016). *Process quality*, on the other hand, is characterized by more dynamic aspects of the classroom, including teacher-child interactions or children's interactions with others in the classroom (Cassidy et al., 2005; Hestenes et al., 2007; La Paro et al., 2012; Pianta et al., 2016). Within early childhood classrooms, high quality teacher-child interactions are characterized by "teachers' sensitivity to individual needs, support for positive behavior, and stimulation of language and cognitive development" (Pianta et al., 2016, p. 119).

Although both *structural* and *process quality* are important contributors to safe, productive early childhood classroom learning environments, research shows that *process quality* is more directly related to child developmental skills (Burchinal et al., 2016; Mashburn et al., 2008; Pianta et al., 2005). It is hypothesized that *process quality* may mediate the relationship between *structural quality* and children's cognitive and social skills. In other words, while it is important to have foundational *structural quality* features in place, such as curriculum and materials, *how* the teacher uses these resources and materials in the classroom more directly influences children's learning (La Paro et al., 2012)

It is clear that teachers' interactions with children directly benefit their early learning and development. However, most research to date on process quality has been conducted in preschool classrooms (Burchinal et al., 2016; Mashburn et al., 2008; Pianta et al., 2005). Few research studies have examined *process quality* in toddler classrooms, or in programs serving diverse, low-income children, which is particularly important because experiences in toddlerhood are predictive of later positive development (National Scientific Council on the Developing Child, 2010; Vandell et al., 2010). Research is

needed to examine measures of process quality in toddler classrooms to inform early interventions that improve quality and that can promote the development of foundational school readiness skills.

Classroom Quality in Toddler Classrooms

Developmentally appropriate high quality teacher-child interactions that support toddlers are similar but also different from those that best support preschool children (Copple & Bredekamp, 2009). During the toddler period, toddlers benefit from consistent responsive interactions and routines that facilitate learning and development, also known as an “experience expectable environment,” which allow toddlers to feel comfortable growing and learning in classrooms (La Paro & Gloeckler, 2016). Identifying high quality classroom interaction processes that support toddler development is critical because the skills foundational to early school success are acquired during the toddler stage. These skills include developing a sense of self, language, motor coordination, learning to interact with others, asserting independence, and developing emotional and behavioral self-regulation, which support successful transition to the social and classroom environment (Brownell & Kopp, 2007; Calkins, 2007; La Paro & Gloeckler, 2016).

One of the most commonly used observational measures to assess the quality of teacher-child interactions in toddler classrooms is the CLASS-T. The CLASS-T includes two domains, Emotional and Behavioral Support, and Engaged Support for Learning (La Paro et al., 2012). The first CLASS-T domain measures emotionally supportive environments that are characterized by a positive emotional climate and teachers who are sensitive to children’s needs. This domain also measures behaviorally supportive environments, which includes the use of behavior guidance strategies by teachers, such as

consistent limits, behavioral feedback through engagement and praise, and supporting positive behavior. The second CLASS-T domain measures the extent to which teacher-child interactions support the development of cognitive and language skills in the context of interactions and relationships through back and forth exchanges of information (Bandel et al., 2014). The domains of the CLASS-T include high quality teacher-child interactions that are identified as some of the most developmentally appropriate interactions that support the development of skills, such as self-regulation and language skills, and allow children to feel comfortable forming secure relationships with teachers and exploring their learning environment (Calkins, 2007; Feldman & Klein, 2003; Gonzalez-Mena & Eyer, 2007; Mortensen & Barnett, 2015).

Recognizing the importance of measuring and supporting high quality teacher-child interactions for toddlers, the CLASS-T is increasingly used in early childhood programs such as EHS and in QRIS programs. However, few research studies have validated the use of the CLASS-T in early childhood programs serving diverse, low-income children. For example, only one factor analytic study has been conducted using the Early Head Start Family and Child Experiences Study (Baby FACES) sample, which comprised a national sample of EHS classrooms (Bandel et al., 2014). Findings provided initial reliability and validity evidence for the two CLASS-T domains of Emotional and Behavioral Support (EBS) and Engaged Support for Learning (ESL) for use in EHS. However, there are several limitations to this study that need to be addressed. First, the factor analyses did not validate the measure in linguistically diverse classrooms, such as classrooms in which teachers' predominant classroom language is Spanish. Second, the Baby FACES study relied upon an older version of the CLASS-T, rather than the

currently available published version used most widely in early childhood programs. More research is needed to extend the validity of the CLASS-T, using the current published version of the measure, and to examine associations between the domain scores of the CLASS-T and toddler developmental skills. Research using the CLASS-T is also needed within other samples of diverse toddlers from low-income families, such as those enrolled in subsidized child care programs and those that are enrolled in classrooms where teachers predominantly use Spanish in the classroom.

Associations between Process Quality and Child Developmental Skills

The relationship between process quality and children's developmental skills has been examined extensively in prekindergarten. Research shows positive associations between domains of process quality and children's social-emotional, language, and cognitive skills in preschool classrooms (Burchinal et al., 2008; Burchinal, Kainz, & Cai, 2011; Howes et al., 2008; Mashburn et al., 2008). However, very few studies have examined the associations between process quality and children's developmental skills in toddler programs (Aikens et al., 2015). The results from toddler studies show that process quality is related to improved cognitive, language, and social-emotional skills for toddlers (La Paro, Williamson, & Hatfield, 2014; Li, Farkas, Duncan, Burchinal, & Vandell, 2013; Loeb, Fuller, Kagan, & Carrol, 2004; McCartney et al., 2007; Mortensen & Barnett, 2015; NICHD, 2002; Ruzek et al., 2014; Vandell et al., 2010). In addition, in longitudinal studies, associations are found later in development, such that process quality experienced by children during infancy and toddlerhood predicts higher cognitive achievement and reduced externalizing behavior at age 15 (Vandell et al., 2010).

Despite these findings, few studies to date have examined the relationship between process quality and children's developmental skills specifically for low-income samples of toddlers. Attending poor quality centers has been shown to be especially problematic for children from low-income backgrounds who face increased risks to their school readiness prior to kindergarten (Watanabe et al., 2011). High process quality during the early childhood period may buffer the negative effect of poverty on academic outcomes (Dearing et al., 2009; McCartney et al., 2007; Sabol & Pianta, 2012; Watanabe et al., 2011), underscoring the importance for children from low-income backgrounds to attend high quality ECE programs. Supporting this idea, research shows that high quality teacher-child interactions are associated with improved language and social-emotional development for toddlers attending EHS (Aikens et al., 2015; Bandel et al., 2014). Because there are so few studies examining the effect of process quality on children's developmental skills in low-income toddler samples, it is important to extend these findings to other samples of toddlers from low-income families, such as those that are enrolled in subsidized child care programs. Extending validity evidence for the use of the CLASS-T to more classrooms serving diverse children in order to better understand these relationships can contribute to national efforts to measure and increase the quality of classroom environments for our nations' youngest learners.

Thresholds of Classroom Process Quality

Despite the important contribution of high quality teacher-child interactions to children's early learning and development, research typically identifies relatively weak linear statistical associations between scores on measures of classroom quality and children's developmental skills. Most studies to date have focused on the linear

relationship between classroom quality and children's developmental skills, which show that increases in increments in quality are associated with equivalent increases of scores on child developmental skills across all levels of quality (Burchinal et al., 2016). In these studies, effect sizes of the regression estimates between indicators of classroom quality and child developmental skills were relatively weak and ranged from .05 and .17, with the majority of effect sizes less than .10 (Burchinal et al., 2011; Burchinal et al., 2016; Keys et al., 2013).

Researchers have speculated that the reason for these modest relationships between quality and children's developmental skills may be due to the use of a linear statistical approach to modeling these associations. In fact, children may need to experience a certain level of quality before quality is positively associated with developmental skills (Burchinal et al., 2016; Le, Schaack, & Setodji, 2015). If there are different relationships between quality and child developmental skills across the range of quality, linear statistical analyses would fail to capture these relationships accurately. The use of nonlinear models is recommended as a more sensitive statistical approach to capture the relationships between classroom quality and children's developmental skills that may be masked with a single linear model (Burchinal et al., 2016).

There are two general nonlinear approaches that are used: (a) an *a priori* approach in which the linear relationship between quality and developmental skills differs in higher and lower quality classrooms based on set cut-points that are imposed beforehand, and (b) an *empirically derived* approach in which thresholds are identified based on plotting classroom quality scores and examining inflection points at which the relationship between quality and developmental skills appears to change. Recent studies in early

childhood programs favor the use of the *a priori* approach because it allows researchers to empirically test whether the cut-points that are used in practice are valid (Burchinal et al., 2016; Hatfield, Burchinal, Pianta, & Sideris, 2016). Research is needed to examine the relationship between classroom quality and child developmental skills in EHS and subsidized child care programs serving low-income children, applying these nonlinear approaches to identify whether there is a threshold of classroom quality at which point children's developmental skills benefit.

Quality thresholds identified in early childhood research. Recent studies have identified thresholds of classroom process quality where quality is more strongly associated with children's developmental skills in preschool. These studies show that the association between classroom process quality and children's developmental skills, including social-emotional, cognitive, and academic skills, is stronger at higher levels of classroom process quality when compared to lower levels of quality (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Burchinal, Vernon-Feagans, Vitiello, Greenberg, & The Family Life Project Key Investigators, 2014; Burchinal et al., 2016; Finch, Johnson, & Phillips, 2015; Hatfield et al., 2016; Le et al., 2015; Weiland, Ulvestad, Sachs, & Yoshikawa, 2013). Longitudinal nonlinear associations with cognitive and social-emotional skills at age 15 were also found for children who experienced a certain threshold of process quality in early childhood programs (Vandell et al., 2010).

Although there have been studies examining thresholds of quality in preschool classrooms, few studies have examined thresholds of classroom quality in toddler classrooms. In one study, researchers found that there were stronger associations between *global* classroom quality, which included both *structural* and *process quality*, as

measured by the Infant/Toddler Environment Rating Scale (ITERS), and cognitive skills at moderate levels of quality when compared to lower levels of quality (Setodji, Le, & Schaack, 2013). In the second study, which used the Baby FACES sample of toddlers attending EHS, researchers identified a threshold of *process quality* on the CLASS-T, at which point quality influenced children's developmental skills more strongly (Aikens et al., 2015). Differential associations were found, such that there was a stronger relationship between the EBS domain scores and social-emotional skills and a stronger relationship between the ESL domain scores and language, problem solving, and social-emotional skills in higher quality classrooms when compared to lower quality classrooms. These initial findings suggest that thresholds of quality, as measured by the CLASS-T domains scores, may exist in the relationship between quality and children's developmental skills; however more research is needed to extend these findings in other ethnically and linguistically diverse low-income toddler samples, such as those attending subsidized child care programs.

Limitations and need for research on thresholds of process quality. Although initial research findings provide evidence to suggest that there may be thresholds of quality in toddler classrooms as measured by CLASS-T scores, there are several limitations to these studies. First, prior studies used *empirically derived* methods to identify thresholds. While these methods are helpful, the cut-points may not be directly generalizable to local programs that are using different, a priori cut-points in their quality improvement efforts. For example, the cut-points that were identified by Aikens and colleagues (2015) are different from the cut-points currently being used by the local Miami-Dade County QRIS, Quality Counts. In order to examine whether the cut-points

being used in practice locally are supported empirically, statistical models need to be tested using an a priori method. Second, the one study that used the CLASS-T to examine thresholds (Aikens et al., 2015), used a pilot measure of the CLASS-T (Pianta, La Paro, & Hamre, 2010). Since then, a revised published version of the CLASS-T was released for public use (La Paro et al., 2012). Additional studies are needed to replicate these findings with the revised version of the CLASS-T as it is now widely used in EHS and subsidized child care programs nationally serving children from diverse low-income backgrounds.

Third, the only study that examined thresholds with the CLASS-T domain scores with a low-income toddler sample used the EHS Baby FACES sample (Aikens et al., 2015). These results may not be generalizable to other populations of toddlers from low-income families who are not enrolled in EHS programs. Community-based programs such as those serving children who receive CCDF funding typically do not receive the same comprehensive set of supports as provided by EHS and are typically rated as lower quality (Johnson, Ryan, & Brooks-Gunn, 2012; Loeb et al., 2004), so the results from previous studies may not be generalizable to these samples. To address these limitations, research is needed to replicate and extend initial research in other samples of ethnically and linguistically diverse toddlers from low-income families, including those enrolled in subsidized child care programs. Overall, research is needed using the revised and published CLASS-T measure applying a priori methods to identify thresholds based on cut-points currently used in practice in diverse, low-income toddler samples that have direct implications for EHS and QRIS programs.

The Current Study

Given the growing policy emphasis on quality improvement based on measures of classroom process quality, such as the CLASS-T, and policy-decisions based on thresholds of classroom quality using CLASS-T scores, empirical research is needed within early childhood programs serving the growing number of ethnically and linguistically diverse populations of children living in poverty. Therefore, the current study examined the validity of the CLASS-T and the linear and non-linear relationships between the CLASS-T domain scores and child developmental skills in an ethnically and linguistically diverse sample of toddlers attending EHS and subsidized child care programs. Additionally, cut-points used in practice by Quality Counts were validated empirically.

Research Objectives and Hypotheses

Objective 1. Extend the reliability and validity of the CLASS-T in an ethnically and linguistically diverse sample of classrooms within Early Head Start and subsidized child care programs. The factor structure, reliability and invariance of the CLASS-T across demographically diverse classrooms, such as those in which teachers' predominant classroom language use was Spanish, was examined. It was hypothesized that the factor structure would be comparable to the previously validated structure in the overall sample, as well as in classrooms in which teachers' predominant classroom language use was reported as either English or Spanish. Specifically, it was expected that there would be two factors identified, Emotional and Behavioral Support (EBS) and Engaged Support for Learning (ESL), which would show high internal consistency reliability.

Objective 2. Examine whether the CLASS-T is associated with the developmental skills, including language and social-emotional skills, of toddlers participating in Early Head Start and subsidized child care programs. This objective extended the concurrent validity of the CLASS-T domain scores in a local sample of toddlers by examining associations between each domain score of the CLASS-T and a set of developmental skills, including language, social-emotional, motor, cognitive, and personal-social skills. It was hypothesized that both the EBS and ESL domains would be positively associated with toddlers' developmental skills, including language and social-emotional skills, such that higher quality teacher-child interactions would be associated with higher child developmental skills, particularly language and social-emotional skills.

Objective 3. Validate the local quality improvement program's cut-points by examining the extent to which the cut-points used in practice align with statistical thresholds that are differentially associated with children's developmental skills. This study extended previous research by examining whether the cut-points that are currently used in practice by the local QRIS, Quality Counts, aligned with statistical thresholds in the relationship between classroom quality and a set of developmental skills, including language, social-emotional, motor, cognitive, and personal-social skills. These relationships were tested with an *a priori* method that used the current Quality Counts cut-points as the *a priori* values followed by a complementary *empirical* method that would confirm the thresholds visually. It was hypothesized that a threshold would be identified at which point a stronger relationship would be found between both domain scores of the CLASS-T, EBS and ESL, and child developmental skills at higher levels of quality. In other words, it was expected that the association between classroom process

quality and children's developmental skills would be stronger at higher levels of process quality when compared to lower levels of process quality. However, because the cut-points that were currently used in practice by Quality Counts differed from the empirically identified thresholds in previous studies (Aikens et al., 2015), it was unknown whether these cut-points would be validated empirically.

CHAPTER 2 METHOD

The current study was conducted within the context of two larger University-Head Start and Child Care research partnership projects: (a) an ACF-funded project called “*An Inquiry Based Approach: Supporting Teachers and Families to Improve Early Development for Infants and Toddlers*” or Early Head Start-University Partnership (EHS-UP), which is an inquiry-based coaching intervention in EHS programs aimed at improving the quality of teacher-child interactions in toddler classrooms and (b) a research partnership between UM and The Children’s Trust called “*Quality Counts Process/Implementation Evaluation,*” which is an evaluation of the local QRIS, Quality Counts. Classrooms were recruited from both projects to increase the variability and range of quality observed.

Participants

All participants were recruited in the spring of 2018. The final sample consisted of 106 teachers (40 from the EHS-UP project and 66 from the Quality Counts project), all of which were female. The majority of teachers were White (59.4%) with others reported as Black or African American (25.5%), Asian (.9%), and other race (3.8%). The remainder (10.4%) of the teachers preferred not to answer this question. With respect to ethnicity, 73.6% of the teachers reported being Hispanic. Of this sample, the majority of teachers (49.1%) reported using predominantly Spanish in the classroom, 40.6% reported using predominantly English in the classroom, and a small percentage of the teachers (9.4%) reported using a mix of both English and Spanish in the classroom. When reporting highest level of education, teachers reported having some college, but no degree (29.2%), a high school diploma or GED (26.4%), a Bachelor’s degree (21.7%), an

Associate's degree (9.5%), a graduate or professional degree (6.6%), or less than a high school diploma (4.7%). The mean years of experience with infants, toddlers, and young children aged 3-5 years in the classroom was 10.75 years ($SD=8.52$, Range= 0-40 years).

The final child sample consisted of 688 toddlers (161 from the EHS-UP project and 527 from the Quality Counts project) between the ages of 11-39 months ($M=22.95$, $SD=7.22$) with 51.6% boys. Of the children whose parents reported their race, the majority were Black or African American (50.5%), White (36.6%), Other, including Arab and Hispanic (6.8%), American Indian or Alaskan Native (1.6%), Mixed (1.3%), and Asian (.6%). The rest of the sample (2.6%) preferred not to answer this question. Of the child sample, 45.3% were Hispanic.

Within the EHS-UP sample, all children met federal criteria to enroll in Early Head Start, meaning that the family annual income was below \$25,100 for a family of four (Federal Register, 2018). Classrooms and children recruited to participate from the Quality Counts project were identified from centers located in a census tract neighborhood at 150% below the poverty line or classrooms with over 50% children receiving subsidized child care.

Procedures

Data for the current project was collected during the 2017-2018 academic year. The current study was approved by the Institutional Review Board at the University of Miami prior to the start of data collection. For the EHS-UP project, the data were collected during Year 1 of the project prior to the start of the intervention so that the effects of the intervention did not affect the results of the current study. In the spring of 2018, trained research assistants from the EHS-UP team and independent observers

contracted by the Early Learning Coalition (ELC) for Quality Counts observed toddler teachers with the CLASS-T. The CLASS-T training occurred over 2 days and was conducted by a certified trainer from Teachstone. All observers took part in reliability testing where they watched and coded five videos and passed reliability, meaning that they coded each dimension of the CLASS-T within one point of the master coder for at least 80% of the codes across all five reliability videos.

Teachers were asked to complete a demographic questionnaire with information on teacher sex, ethnicity, teaching experience, teacher education, predominant language used in the classroom, and information on the demographic composition of each classroom, including age, sex, ethnicity, and DLL status for each child in the class. Additionally, during the spring of 2018, teachers completed the Brief Infant-Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2002) to assess the social-emotional development of the consented children in each classroom and the Ages and Stages Questionnaires (ASQ-3; Squires, Twombly, Bricker, & Potter, 2009) to assess various developmental domains, including communication and problem solving skills. Across both projects, all of the teachers that filled out the demographic questionnaire and the assessments on children's developmental skills received \$25 as an incentive for participating.

Measures

Classroom quality.

Process quality. Process quality was assessed using the Classroom Assessment Scoring System-Toddler (CLASS-T; La Paro et al., 2012). The CLASS-T is an observational tool that measures the quality of teacher-child interactions in toddler classrooms. The CLASS-T includes two domains, each made up of dimensions:

Emotional and Behavioral Support (EBS), which is comprised of Positive Climate, Negative Climate, Teacher Sensitivity, Regard for Child Perspectives, Behavior Guidance; and Engaged Support for Learning (ESL), which is comprised of Facilitation of Learning and Development, Quality of Feedback, and Language Modeling.

Observers rate each dimension on a 7-point Likert scale, where 1 indicates the lowest quality interactions and 7 indicates the highest quality interactions. These ratings are based on observed interactions between the teacher and children that occur in the classroom during a 20-minute cycle. Each classroom is observed for three cycles and the dimension scores are averaged across the cycles to indicate a classroom score for each dimension.

The EBS domain is characterized by positive climate and emotional connection between the teacher and child, responsiveness and awareness of the teacher, the teacher's recognition of a child's interests and autonomy, and the teacher's use of effective methods to guide children's behavior. The ESL domain is characterized by the teacher's ability to facilitate children's learning and development and teacher's use of language to facilitate children's language development. Both domains show high internal consistency, with Cronbach's alphas of .83 for EBS and .75 for ESL in the current study sample. Criterion-related and convergent validity were established based on relationships with other valid measures of classroom quality and children's developmental skills (La Paro et al., 2012). The CLASS-T has been validated for use with toddlers from low-income families using the EHS national Baby FACES sample (Bandel et al., 2014).

Child developmental skills.

Social-emotional skills. Children's social-emotional skills were assessed using the Brief Infant-Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2002). The BITSEA is a screener that is based on caregiver report. The BITSEA includes 42 items that are used to identify social-emotional and behavioral problems and competencies for toddlers between the ages of 12 and 36 months. The screener is a shorter version of the parent measure, the Infant Toddler Social Emotional Assessment (ITSEA; Carter & Briggs-Gowan, 2006). Caregivers rate each item of the BITSEA on a three-point Likert scale, including "rarely," "sometimes," or "often." The items comprise two factors, the Problem and Competency scales, which showed adequate internal consistency within the current study sample, with Cronbach's alphas of .83 and .79, respectively. Criterion-related and discriminant validity of the BITSEA subscales scores were established through relationships with other validated measures of social-emotional development, and based on the measure's sensitivity and specificity of identification of children rated above the clinical cutoff for behavior problems (Briggs-Gowan, Carter, Irwin, Wachtel, & Cicchetti, 2004). The BITSEA was also validated for use with toddlers from low-income families (Hungerford, Garcia, & Bagner, 2015).

Developmental skills. Children's developmental skills were assessed using the Ages and Stages Questionnaires (ASQ-3; Squires et al., 2009). The ASQ-3 is a screener based on caregiver report with 30 items that are used to assess children's developmental skills in the following domains: Communication, Gross Motor, Fine Motor, Problem Solving, and Personal-social skills. There are different versions of the ASQ-3 that include age appropriate developmental indicators for children between the ages of 1 month to 5 ½

years old. The appropriate version to administer is determined based on the child's birth date and the date of assessment. Caregivers report either "yes", "sometimes", or "not yet" to determine if a child was meeting the developmental domains. All of the domains of the ASQ-3 show high internal consistency reliability (Communication, $\alpha=.83$; Gross Motor, $\alpha=.74$; Fine Motor, $\alpha=.81$; Problem Solving, $\alpha=.77$; and Personal-social, $\alpha=.76$; Squires et al., 2009). The ASQ-3 was validated for use with diverse populations of children and has shown high concurrent validity and sensitivity and specificity as a screening tool (Squires et al., 2009).

Data Analytic Plan

Objective 1: Extend the reliability and validity of the CLASS-T in an ethnically and linguistically diverse sample of classrooms within Early Head Start and subsidized child care programs. To examine the construct validity of the CLASS-T domain scores in this sample of toddlers, confirmatory factor analyses (CFA) were used. Internal consistency reliability was calculated using Cronbach's alpha. In order to determine whether the same factors were identified across classrooms in which teachers reported using predominantly English or Spanish in the classrooms, factorial invariance was examined.

Confirmatory factor analyses. A series of CFAs were examined to determine the fit of the published two-factor structure of the CLASS-T. These analyses were conducted in Mplus Version 7 (Muthen & Muthen, 1998-2011). To assess the fit of the overall model to these data, the following fit statistics were examined; the Comparative Fit Index (CFI > .95 representing good fit and CFI > .90 representing adequate fit; Bentler, 1990; Hu & Bentler, 1999), the Tucker-Lewis Reliability Index (TLI) greater than .90

considered adequate fit (Bentler, 1990; Hu & Bentler, 1999), the root mean square error of approximation (RMSEA < .06 representing close fit and RMSEA < .08 representing reasonable fit; Browne & Cudeck, 1993; Steiger, 1990), and the Standardized Root Mean Square Residual (SRMR) with values <.08 considered adequate fit (Hu & Bentler, 1999). Additionally, the significance and salience (whether the factor loadings were greater than .40; Gorsuch, 1983), of the factor loadings were examined to determine if each dimension loaded significantly on each of the published factors.

Internal consistency reliability. To assess the internal consistency reliability of each CLASS-T factor, Cronbach's alphas were calculated. Values of 0.70 or higher indicated adequate internal consistency reliability (Gorsuch, 1983).

Factorial invariance. Due to the linguistic diversity of the classrooms in this sample, factorial invariance of the CLASS-T across subgroups of classrooms based on the predominant language teachers reported using in the classroom was tested. Teachers were asked to report what languages they used in the classroom and how often they used each language. Based on these responses, classrooms were classified as either majority English, majority Spanish, or a mix of both. The majority of classrooms (over 90%) were characterized as either majority English or majority Spanish, and they were categorized into two groups of classrooms that were used to examine factorial invariance. Configural invariance was examined using Mplus Version 7 (Muthen & Muthen, 1998-2011). These analyses tested whether the same factor structure was invariant across classrooms regardless of whether teachers used predominantly English versus predominantly Spanish. To examine configural invariance, the established CLASS-T factor structure was simultaneously fitted to the covariance structure of two groups: classrooms where

teachers reported using predominantly English and classrooms where teachers reported using predominantly Spanish. For the invariance analyses, the same fit statistics as provided for the CFA above were examined to determine adequate model fit.

Objective 2: Examine whether the CLASS-T is associated with the developmental skills, including language and social-emotional skills, of toddlers participating in Early Head Start and subsidized child care programs. To extend the concurrent validity of the CLASS-T, the linear relationships between the CLASS-T factors and child developmental skills were examined using multilevel regression analyses in R (R core team, 2016). Multilevel modeling (Raudenbush & Bryk, 2002) was used because the main predictors of interest, the domains of the CLASS-T, were measured at the classroom-level (level 2) and child developmental skills were measured at the child-level (level 1). In the first unconditional model, intraclass correlations were examined to identify whether there was variability across child developmental skills to be predicted at level 2 (classroom level). Then, a series of two-level models were estimated where the classroom was treated as the level 2 cluster unit with children nested within classrooms. Child-level demographic variables, including child age and sex were included as covariates at level 1. Separate models were estimated for each developmental domain score, including the domains of the ASQ-3 (Communication, Gross Motor, Fine Motor, Problem Solving, and Personal-social) and the social-emotional domain scores on the BITSEA (Social-emotional Competency and Problems). In each of these models, the two CLASS-T factors (EBS and ESL), were entered as level 2 predictors of the classroom-level variance in each child developmental skill.

The equations for the multilevel models are:

$$\text{Level 1: } ChildSkill_{ij} = B_{0j} + B_{1j}age_{ij} + B_{2j}sex_{ij} + e_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}EBS_j + \gamma_{02}ESL_j + u_{0j}$$

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

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

In the level 1 equation above, $ChildSkill_{ij}$ represented each child developmental skill score for a child (i) who was in classroom (j), including the Competency and Problem subscales of the BITSEA and the Communication, Gross Motor, Fine Motor, Problem Solving, and Personal-social domains of the ASQ-3, predicted by classroom quality. $ChildSkill_{ij}$ was a function of the intercept, B_{0j} (the developmental skill mean score for children in each classroom), the fixed effects associated with the covariates age and sex, respectively (B_{1j} , B_{2j}), and e_{ij} , which is the level 1 residual that represents within classroom variability. In the level 2 equations, the intercept, β_{0j} was a function of the grand mean of the child developmental skills, γ_{00} , the fixed effects for each process quality domain of the CLASS-T, EBS (γ_{01}) and ESL (γ_{02}), and the level 2 residual, u_{0j} , which represents the variability between classrooms. Level 2 equations for each of the child-level predictors represented the fixed effects of each child-level covariate (B_{1j} , B_{2j}) and are a function of the estimated mean differences in the child developmental skills across classrooms per unit change in the child level predictors (γ_{10} , γ_{20}).

Objective 3: Validate the local quality improvement program's cut-points by examining the extent to which the cut-points used in practice align with statistical thresholds that are differentially associated with children's developmental skills. To validate the cut-points used in practice locally, two separate, but complementary statistical analyses were used to examine associations between scores on the CLASS-T domains and children's developmental skills. The first *a priori* approach imposed the cut-

point scores from the CLASS-T domains used by the local QRIS, Quality Counts. Using the a priori approach, a non-linear statistical model examined whether the nature of associations between the CLASS-T and child developmental skills changed at that threshold score. The second *empirical* approach allowed for exploratory statistical estimation of the regression line, plotting graphically the non-linear associations between the CLASS-T and child developmental skills. Findings from the two approaches were compared to determine if the thresholds identified in the a priori approach were consistent with the cut-points apparent when the data were examined visually using the empirical approach.

A priori approach. To examine whether the cut-points that were being used in practice by Quality Counts were valid, piecewise regression or “spline” multilevel analyses were conducted in SAS 9.4 following the approach used by Burchinal et al. (2016) and Hatfield et al. (2016). This *a priori* method uses a spline technique to estimate piecewise regression multilevel models (Greene, 2000; Marsh & Cormier, 2002). In this spline method, local thresholds of classroom quality, as measured by the CLASS-T and determined by Quality Counts, were imposed in the models. The cut-points that were used were chosen based on the cut-points that contributed to the different rating levels within Quality Counts at the time of the present study. The CLASS-T domain scores that were chosen as cut-points were those corresponding to level 4 out of 5 levels, which was considered high quality. A level 4 corresponded to CLASS-T scores of 5.50 or higher on the EBS domain and a 3.50 or higher on the ESL domain.

The piecewise regression multilevel analyses estimated separate splines, or slopes, for higher and lower quality programs based on these cut-points, which allowed

the intercept and slope to vary in higher and lower quality ranges across all child developmental skills. Separate slopes were estimated for the relationship between each CLASS-T factor and child developmental skill for lower and higher quality classrooms. The two slopes for the relationship between the CLASS-T and each child developmental skill in high and low quality classrooms were compared, using an interaction term, to examine if the relationships between the CLASS-T and child developmental skills were statistically different in higher and lower quality classrooms. If the interaction term was significant, this suggested that the slopes in higher versus lower quality classrooms were significantly different from each other; and thus provided empirical support for the a priori cut-point on the CLASS-T as a threshold at which quality was differentially associated with child developmental skills (Burchinal et al., 2014). For example, classrooms rated as 5.50 and higher on the EBS domain were considered higher quality classrooms and classrooms rated lower than 5.50 were considered lower quality classrooms. The relationship between EBS and each child developmental skill was estimated separately within each group, in the classrooms rated over 5.50, or the high quality classrooms, and the classrooms rated under 5.50, or the low quality classrooms. The slopes were then examined to determine if they were statistically significantly different from one another and if they were, this would suggest that there was a threshold aligned with the a priori cut-point, in this example, 5.50.

Separate piecewise regression multilevel models were used to predict each set of the child developmental skills, which were centered at the grand mean: teacher-rated social-emotional development (Competency and Problem subscales of the BITSEA), and teacher-rated developmental domains (Communication, Gross Motor, Fine Motor,

Problem Solving, and Personal-social domains of the ASQ-3). Spline knots were used to set the a priori cut-points to indicate where the slope was allowed to change. The multilevel models accounted for the nesting of children in classrooms at level 2. Two-level, random-intercepts Hierarchical Linear Models (HLM) were used. In the HLM analyses, a residual term was included to take into account the clustering of children within classrooms and an independent residual term was included to account for the error in the individual children's scores. All of the models controlled for child characteristics, including child age and sex.

The equation for the piecewise regression multilevel model was:

$$ChildSkill_{ij} = B_0 + B_1 Quality_j + B_2 Quality_j * HighQualityClassroom_j + B_3 Age_{ij} + B_4 Sex_{ij} + e_{ij} + u_j$$

HighQualityClassroom_j was a dummy variable that included a value of 1 if the classroom quality for that classroom was at or above the high-quality cut-point (5.50 for EBS and 3.50 for ESL). *Quality_j* was centered at the first value of quality in the range of quality so that the interaction was more interpretable, as recommended by Greene (2000). In the equation above, *ChildSkill_{ij}* represented each child developmental skill, including the Competency and Problem subscales of the BITSEA and the Communication, Gross Motor, Fine Motor, Problem Solving, and Personal-social domains of the ASQ-3, predicted by classroom quality. B_0 was the overall intercept, B_1 corresponded to the slope in lower quality classrooms, B_2 corresponded to the difference between the slopes in higher and lower quality classrooms (with $B_1 + B_2$ corresponding to the estimated slope for the higher quality classrooms), which included an interaction between the quality indicator and the dummy variable identifying high quality classrooms as determined by

the cut-points. $B_3 - B_4$ were the coefficients for the covariates that were included in the model. This model also accounted for nesting of children within classrooms by including the level 1 residual, e_{ij} , which represented within classroom variability, and the level 2 residual, u_j , which represented the variability between classrooms. If B_2 was statistically significant, which was the difference between the slopes of the high and low quality classrooms, then it was determined that there was a threshold at the a priori cut-point.

Empirical approach. Following the piecewise regression multilevel analyses, an empirical approach was used to confirm the findings from the piecewise regression multilevel models. These procedures followed methods used by Burchinal and colleagues (2016) and Hatfield and colleagues (2016). A b-spline model was used because it is a flexible approach for modeling the association between the domains of the CLASS-T and children's developmental skills. The goal of the b-spline model is to find a function that either interpolates points or fits a smooth curve through them (Burchinal et al., 2016). This model is used to model data in which the relationship is nonlinear and the parametric form is not obvious. The nonlinear terms were tested to determine if they were significantly different from zero. If so, plots were used to examine whether there was evidence visually of thresholds in the relationship between the CLASS-T and child developmental skills and whether these thresholds supported the findings from the piecewise regression multilevel models.

The model allowed for up to three knots (allowing for up to three thresholds), as has been examined in previous research (Burchinal et al., 2016; Hatfield et al., 2016). The additional knots allowed for more flexibility when examining the shape of the line. The equation that was used to examine this relationship was:

$$F(x) = B_0 + B_1x + B_2x^2 + B_3x^3 + B_4(x-k_1)^3 + B_5(x-k_2)^3 + B_6(x-k_3)^3$$

In the above equation, x was the quality indicator being used and k_1-k_3 were the knots that were defined by the 25th, 50th, and 75th percentiles in that measure's distribution. A separate curve was fit between each of the knots, for the four quartiles examined, which allowed for greater flexibility in the shape of the overall function. The model would add a "penalty" that forced each piece to join the others at each knot (Burchinal et al., 2016).

CHAPTER 3 RESULTS

Descriptive Statistics

To examine whether the data were normally distributed, the variables included in the study were analyzed for outliers, homoscedasticity, and kurtosis. None of the assumptions of normality were violated. Table 1 includes descriptive statistics for all of the measures in the study, including means and standard deviations for this sample. Additionally, Table 2 includes bivariate correlations among the measures included in the study.

Objective 1: Construct Validity of the CLASS-T

Confirmatory factor analyses. Confirmatory factor analyses were used to examine the factor structure of the CLASS-T. Both a one-factor model, where all dimensions of the CLASS-T loaded onto one factor of teacher-child interaction quality, and a two-factor model, in which dimensions were loaded onto two domains of teacher-child interaction quality, Emotional and Behavioral Support (EBS) and Engaged Support for Learning (ESL), were examined. Both the one-factor model ($\chi^2 = 45.09$, $p < 0.001$; CFI = 0.93, TLI = 0.87, RMSEA = 0.13, SRMR = 0.06) and the two-factor model ($\chi^2 = 31.12$, $p = 0.009$; CFI = 0.96, TLI = 0.92, RMSEA = 0.10, SRMR = 0.07) fit the data adequately, although the two-factor model fit the data better based on the established fit criteria. To identify whether the one-factor or two-factor model would fit the data best, the Satorra Bentler χ^2 difference test was examined. The results were significant, indicating that the two-factor model fit these data best ($\chi^2_{\text{diff}}(1) = 13.96$, $p < 0.001$). See Figure 1 for the factor loadings of the final two-factor model in the overall sample. All of the factor loadings loaded significantly onto each of the two factors, and the loadings

were salient based on the established criteria, as almost all of the factor loadings were over .40 (Gorsuch, 1983). The only factor loading that was not over .40 was Negative Climate. However, the dimension loaded significantly to the EBS domain. Given prior research and theory on the CLASS-T (Bandel et al., 2014) and slight variations that might emerge as factor structures are dependent on the local sample, the two-factor CLASS-T structure with the Negative Climate dimension was retained.

Internal consistency reliability. Cronbach's alphas were calculated to examine the internal consistency reliability of each of the two CLASS-T factors. The Cronbach's alpha for EBS was .83 and the Cronbach's alpha for ESL was .75, both indicating good internal consistency reliability (Gorsuch, 1983).

Measurement invariance. Factorial invariance of the two-factor structure of the CLASS-T was examined using configural invariance to examine whether the same factor structure was identified across classrooms where teachers used predominantly English versus Spanish. In the majority of classrooms in the sample, 49.1% ($n=52$) of teachers reported using predominantly Spanish during the day, while in 40.6% ($n=43$) of classrooms teachers reported using predominantly English during the day. A small percentage of the classrooms (9.4%; $n=10$) were characterized by the use of a mix of both English and Spanish. These classrooms were not included in the analyses.

Configural invariance was examined through multiple group confirmatory factor analysis comparing the factor structure of the CLASS-T in classrooms where teachers spoke predominantly English versus Spanish. These analyses applied the same factor structure to both groups of classrooms allowing all the parameters to be estimated freely (Horn & McArdle, 1992). Results showed that the model fit the data adequately, although

not very well ($\chi^2 = 99.06$, $p < 0.001$; CFI = 0.86, TLI = 0.84, RMSEA = 0.15, SRMR = 0.19). It is possible that the model did not fit these data very well because the analyses were underpowered. Multiple group CFAs typically require a sample of 200 or more participants (Myers, Ahn, & Jin, 2011) and the current sample only included between 40-50 classrooms per language group.

Even though the two-factor model did not fit these data very well, the results showed that all of the factor loadings of the CLASS-T dimensions for the predominantly English language classrooms were significant and salient, as the factor loadings were over .40, and most of the factor loadings for the predominantly Spanish language classrooms were significant and salient. See Figures 2-3 for the factor loadings for each group. The only factor loading that was not significant or salient was the Negative Climate dimension on the EBS domain in the Spanish language classrooms. Given that these analyses were underpowered, findings were interpreted as providing preliminary evidence for the configural invariance of the two-factor structure across the two groups of classrooms. The two-factor solution with the Negative Climate dimension was retained for use throughout the remainder of the analyses based on prior evidence in other low-income toddler samples (Bandel et al., 2014) and because the two-factor structure was supported in the overall sample.

Objective 2: Linear Relationship Between the CLASS-T and Child Developmental Skills

Unconditional models. To examine the linear relationship between the CLASS-T and child developmental skills, a series of multilevel models were tested. Separate models were tested for each of the developmental domains that were assessed using the BITSEA and ASQ-3. The first set of models examined were unconditional models to

determine the variance in each of the child developmental skills at Level 1, the variability attributed to differences between children within classrooms, and at Level 2, the variability attributed to differences between classrooms. The ICCs from the unconditional models showed that a significant proportion of the variance in both social-emotional problems and competence using the BITSEA (63% and 61%, respectively) was attributable to child-level differences while the remaining variance (37% and 39%) was attributed to differences between classrooms. The ICCs for the five domains of the ASQ-3 from the unconditional models showed that a large proportion of the variance in all domains (ranging from 72% - 98%) was attributable to child-level differences and the remaining variance (ranging from 2% - 28%) was attributed to differences between classrooms.

Social-emotional problems. The final multilevel model results examining the relationship between the *EBS* domain score and BITSEA *social-emotional problem* score, while controlling for child age and sex, showed that there was a significant negative relationship between EBS and social-emotional problems ($B = -1.40$, $SE = 3.84$, $p = .035$). Classrooms with higher EBS scores were associated with lower classroom levels of child social-emotional problems. Similarly, there was also a significant negative relationship between the *ESL* domain score and *social-emotional problems* ($B = -1.24$, $SE = 0.54$, $p = .024$), controlling for child age and sex.

Social-emotional competence. The results for the relationship between *EBS* and BITSEA *social-emotional competence*, while controlling for child age and sex, were not significant ($B = 0.25$, $SE = 0.38$, $p = .515$). The results for the relationship between *ESL* and *social-emotional competence*, while controlling for child age and sex, showed that

there was a trend level positive relationship between ESL and social-emotional competence ($B = 0.60$, $SE = 0.31$, $p = .057$), indicating that classrooms rated higher on ESL were associated with higher classroom levels of child social-emotional competence.

None of the linear relationships between either of the CLASS-T domains and the ASQ domains were statistically significant. See Tables 3-4 for the results of all of the multilevel models examined, including the relationships between each child developmental skill and the child level covariates, age and sex.

Objective 3: Thresholds in the Relationship Between the CLASS-T and Child Developmental Skills

A priori approach. In order to examine whether the a priori cut-points used within Quality Counts aligned with thresholds in the relationship between the CLASS-T and child developmental skills, piecewise regression multilevel analyses were examined, controlling for child age and sex. Results are presented in Table 5.

CLASS-T EBS cut-point and social-emotional problems. The results of the piecewise regression multilevel analyses between *EBS* and BITSEA *social-emotional problems* showed that there was a significant negative relationship between EBS and social-emotional problems in classrooms rated as a 5.50 or higher on EBS ($B = -2.56$, $SE = 0.90$, $p = .006$), such that higher EBS was associated with lower classroom levels of social-emotional problems. For classrooms rated as lower than 5.50, there was no significant relationship between EBS and social-emotional problems ($B = 0.02$, $SE = 0.15$, $p = .892$). The two slopes were found to be significantly different from one another, indicating that there was in fact a threshold at the cut-point of 5.50 on EBS. Overall, findings suggested there was a stronger, negative relationship between scores on the EBS

domain of the CLASS-T and social-emotional problems at higher levels of quality when compared to lower levels of quality.

CLASS-T ESL cut-point and social-emotional problems. Thresholds were identified in the relationship between *ESL* and *social-emotional problems*. Results showed that there was no significant relationship between ESL and social-emotional problems at higher levels of quality, above a 3.50 on the ESL domain ($B = -1.09, SE = 0.86, p = .208$). However, there was a significant positive relationship between ESL and social-emotional problems at lower levels of quality, in classrooms rated lower than 3.50 on the ESL domain of the CLASS-T ($B = 0.65, SE = 0.27, p = .018$). This suggested that higher ESL scores, in classrooms rated lower than 3.50, were associated with higher classroom levels of social-emotional problems. The slopes were significantly different from one another, suggesting that there was a threshold aligned with a cut-point of 3.50 on the ESL domain of the CLASS-T. Overall, findings suggest that there was a stronger, positive relationship between ESL and social-emotional problems at lower levels of quality that was no longer significant at higher levels of quality.

CLASS-T EBS cut-point and social-emotional competence. Results from the piecewise models also supported a threshold in the relationship between *EBS* and *social-emotional competence*. There was a significant positive relationship between EBS and social-emotional competence in classrooms rated as 5.50 or higher on the EBS domain ($B = 2.00, SE = 0.54, p < .001$), indicating that EBS scores were associated with higher classroom levels of social-emotional competence at higher levels of quality. However, there was no significant relationship between EBS scores and social-emotional competence at lower levels of quality, when classrooms were rated lower than 5.50 on

EBS ($B = 0.15$, $SE = 0.09$, $p = .087$). The two slopes were significantly different from one another, indicating that there was a threshold at the cut-point of 5.50 on EBS. Overall, findings suggested there was a stronger, positive relationship between the CLASS-T and social-emotional competence at higher levels of quality when compared to classrooms at lower levels of quality.

CLASS-T ESL cut-point and social-emotional competence. Similar threshold results were found for the piecewise models examining the relationship between *ESL* and *social-emotional competence*. There was a positive relationship between ESL and social-emotional competence at higher levels of quality, when classrooms were rated higher than 3.50 on ESL ($B = 1.46$, $SE = 0.51$, $p = .005$). However, there was no significant relationship between ESL and social-emotional competence at lower levels of quality, below 3.50 on ESL ($B = -0.03$, $SE = 0.16$, $p = .848$). These slopes were significantly different from one another suggesting that there was a threshold that aligned with the a priori cut-point of 3.50 in the relationship between ESL and social-emotional competence.

There were no thresholds identified in the relationship between the CLASS-T and any of the ASQ-3 domain scores.

Empirical approach. A complimentary analysis, called a b-spline model, was examined to support what was found with the a priori approach. Because there were no thresholds identified in the relationship between the CLASS-T and any of the ASQ-3 developmental domains using the a priori approach, the empirical approach was not applied to those data. For the BITSEA, empirical models were designed to allow for up to 3 thresholds in the association between each CLASS-T domain and BITSEA subscale

scores of child social-emotional problems and competence, while controlling for child age and sex. Overall, all of the analyses using the empirical approach, that graphically plotted the nonlinear relationships, replicated the findings from the a priori analyses. Each model suggested that there were significant nonlinear relationships between each of the CLASS-T domains and BITSEA social-emotional problems and competence, based on the significance of the F tests and the visual plots. Table 6 includes the results from the empirical b-spline analyses for the BITSEA models. Figures 4 through 7 include the plotted data in the nonlinear relationships found between each CLASS-T domain and the BITSEA domains.

EBS and social-emotional problems. The results for models examining the relationship between *EBS* and *social-emotional problems* were similar to what was found with the a priori approach. The empirical models suggested that there was no clear relationship between EBS and social-emotional problems at lower levels of quality, indicated by a flat slope at lower levels of EBS. As scores on the EBS domain approached higher levels of quality, the plot of the slope of the line showed a steeper downward slope, suggesting a negative relationship between EBS and social-emotional problems. The visual plot was similar to the non-linear relationships found using the a priori approach. Both the a priori and empirical approaches showed that there was a stronger negative relationship between EBS and social-emotional problems at higher levels of quality when compared to lower levels of quality. See Figure 4 for the visual plot of the relationship between EBS and social-emotional problems.

ESL and social-emotional problems. The results for models examining the relationship between *ESL* and *social-emotional problems* suggested that there was a

positive relationship between ESL and social-emotional problems at lower levels of quality, as evidenced by a strong upward slope in the visual plot, at lower levels of ESL. However, as ESL scores increased, that slope seemed to level out to a flat slope. As ESL scores increased further, there seemed to be a second inflection point where there was a downward trend in the slope, indicating a negative relationship between ESL and social-emotional problems at higher levels of quality. The visual plots were similar to the nonlinear relationships found with the a priori approach, where there was a positive relationship between ESL and social-emotional problems at lower levels of quality that was no longer significant at higher levels of quality. However, the downward slope at higher levels of ESL was not identified in the a priori analyses. These findings from the empirical approach add to the a priori findings and suggest that there may be an additional threshold, at which there is a negative relationship between ESL and social-emotional problems at higher levels of quality. See Figure 5 for the visual plot of the relationship between ESL and social-emotional problems.

EBS and social-emotional competence. The results for the models examining the relationship between *EBS* and BITSEA *social-emotional competence* were similar to what was found using the a priori approach. The empirical models suggested no clear relationship at lower levels of quality, as evidenced by a flat slope at lower levels of EBS. As EBS scores increased to higher levels of quality, the slope seemed to change and there was a strong upward slope, indicating a positive relationship between EBS and social-emotional competence at higher levels of EBS. These visual plots followed a similar trend observed for the thresholds found in the a priori analyses, in which there was a stronger positive relationship between EBS and social-emotional competence at higher

levels of quality when compared to lower levels of quality. See Figure 6 for the visual plot of the relationship between EBS and social-emotional competence.

ESL and social-emotional competence. The results from the models examining the relationship between *ESL* and *social-emotional competence* suggested that at lower levels of EBS, there seemed to be a slight downward slope, indicating a negative relationship between ESL and social-emotional competence. As ESL scores increased, the slope seemed to level off, as evidenced by a flat slope, and as ESL scores continued to increase, there seemed to be a second inflection point where the slope turned into an upward slope, indicating a positive relationship with social-emotional competence at higher levels of ESL. The visual trend suggested a positive relationship between ESL and social-emotional competence at higher levels of ESL, as was found with the a priori approach. The visual plots from the empirical approach were similar to the a priori approach except that the a priori approach did not identify a significant negative relationship between ESL and social-emotional competence at lower levels of quality. Findings from the empirical approach add to the findings from the a priori approach and suggest that there might be another threshold in the relationship between ESL and social-emotional competence at which there is a negative relationship between ESL and social-emotional competence at lower levels of quality, which levels off and becomes positive at higher levels of ESL. See Figure 7 for the visual plot of the relationship between ESL and social-emotional competence.

CHAPTER 4 DISCUSSION

This study addressed some of the gaps in the early childhood field by extending the construct validity evidence for the use of the domain scores of the CLASS-T in an ethnically and linguistically diverse sample of toddlers within both EHS and subsidized child care programs. Evidence for the linear and non-linear relationship between the CLASS-T and child social-emotional problems and competence was found. In addition, empirical support for the cut-points being used within the local QRIS, Quality Counts, was found such that scores on the CLASS-T that were associated with the Quality Counts cut-points for high quality corresponded to statistical thresholds in the relationship between the CLASS-T and child social-emotional problems and competence. Overall, results extended the validity evidence for the use of CLASS-T scores to measure the quality of teacher-child interactions in the classroom with a diverse sample of toddlers enrolled in EHS and subsidized child care classrooms.

Construct Validity of the CLASS-T

As hypothesized, the results from the Confirmatory Factor Analyses (CFA) supported the published two-factor structure of the CLASS-T (La Paro et al., 2012). Two domains, Emotional and Behavioral Support (EBS) and Engaged Support for Learning (ESL), emerged as reliable dimensions of teacher-child interaction quality in this sample of diverse toddler classrooms. The EBS domain measures the extent to which to which teacher-child interactions within the classroom are characterized by a positive emotional climate, teachers that are sensitive to children's needs, and a behaviorally supportive environment, including the use of guidance strategies by teachers, such as consistent limits, behavioral feedback through engagement and praise, and supporting positive

behavior. The ESL domain measures the extent to which teacher-child interactions support the development of cognitive and language skills in the context of interactions and relationships through back and forth exchanges of information.

The two-factor structure of the CLASS-T identified in this sample comports with initial research conducted with the CLASS-T in the national EHS, Baby FACES sample (Bandel et al., 2014). Results extend prior research by providing additional validity evidence for the use of the two-factor structure of the CLASS-T in diverse samples of low-income toddlers, including those attending subsidized child care programs. Extending validity evidence for the use of the two-factor structure of the CLASS-T in diverse low-income samples is important as the CLASS-T is used increasingly in community-based programs, such as those serving children who receive CCDF funding. Community-based child care programs typically do not receive the same level of comprehensive supports as EHS and are usually rated as lower quality (Johnson et al., 2012; Loeb et al., 2004). This study provided additional support for the use of measures such as the CLASS-T within community-based child care programs so that teacher-child interaction quality can be measured, monitored, and greater support provided to enhance classroom quality within these programs.

To provide additional support for the use of the two-factor structure in linguistically diverse classrooms, measurement invariance was tested based on teacher predominant classroom language use. The results provided preliminary support that the published two-factor structure of the CLASS-T was valid in classrooms where teachers used predominantly English and predominantly Spanish, as was hypothesized. However, one dimension of the CLASS-T, Negative Climate, fit the two-factor structure in the

English language classrooms but did not fit in the Spanish language classrooms. This may indicate that these analyses are capturing more than just linguistic differences, but also cultural differences attributed to the Hispanic background of the teachers in the classrooms. Negative Climate was very rarely, if at all, observed in the Spanish language classrooms and it is likely because of cultural expectations in classrooms where teachers are Hispanic, and many children enrolled are Hispanic. For example, the cultural value of *respeto* emphasizes obedience and respect for authority figures, such as teachers (Calzada, Fernandez, & Cortes, 2010; Gonzales-Ramos, Zayas, & Cohen, 1998), supporting the idea that children are less likely to act out in the classroom, which in turn would result in less negative affect and punitive control from the teacher. In addition, Hispanic cultural values emphasize creating warm and supportive relationships with children (Guilamo-Ramos et al., 2007), suggesting that teachers would be less likely to use punitive control and show negativity towards children, which would also result in lower scores on the Negative Climate dimension in Spanish language classrooms.

Although the model did not fit these data very well, likely because of the small sample size, these findings contribute to the literature as there has been no research conducted to date examining the measurement invariance of the CLASS-T in linguistically diverse classrooms. Examining measurement invariance across linguistically diverse classrooms is particularly important because the linguistic diversity of the early childhood teacher work force and children in EHS and subsidized childcare programs, has increased and is projected to increase in the coming years (Frey, 2018; Jiang, Granja, & Koball, 2017b; Whitebook, McLean, Austin, & Edwards, 2018). Findings provide support that in classrooms where teachers use predominantly Spanish,

the quality of teacher-child interactions can be measured by the two CLASS-T domains, one that is characterized by both an emotionally and behaviorally supportive environment and one that is characterized by an environment that supports both language and cognitive skills. Future research is needed in larger samples to examine the measurement invariance in these two demographic subgroups of classrooms to identify whether the two-factor structure fits best in these diverse classrooms or whether another factor structure would fit better.

Linear Relationship Between the CLASS-T and Child Developmental Skills

The linear relationship between the two CLASS-T domains and child developmental skills was examined using multilevel regression analyses. The results from these analyses suggested that there was a linear relationship between the quality of teacher-child interactions and social-emotional problems, as well as social-emotional competence at a trend level, as was hypothesized. Higher classroom scores on the EBS and ESL domains were negatively associated with social-emotional problems in the classroom. Additionally, ESL scores were positively associated (at a trend level) with social-emotional competence in the classroom. These results are similar to those in the Baby FACES sample, in which they found that high quality teacher-child interactions, as measured by the CLASS-T domain scores, were associated with lower social-emotional problems and higher social-emotional competence using the BITSEA (Aikens et al., 2015; Bandel et al., 2014).

These findings extend the very limited research on the relationships between the CLASS-T and children's developmental skills by examining these relationships, not only in EHS, but also in subsidized child care programs. As expected, findings suggest that

higher quality teacher-child interactions in the classroom are associated with positive social-emotional development in low-income toddler samples. In classrooms where teachers build positive relationships with children, are sensitive to their needs, and where they are proactive in supporting positive behavior, these interactions may model positive interactions with others and support children in regulating their behavior, which may contribute to lower social-emotional problems and higher social-emotional competence (La Paro et al., 2014). Additionally, teacher-child interactions that support both cognitive skills, such as problem solving, and language skills provide children with constructive ways of expressing their emotions and needs, and solving problems, which contributes to more positive social-emotional development (Mortensen & Barnett, 2015). In summary, the associations that were found between the quality of teacher-child interactions and social-emotional skills in toddler classrooms highlight the importance of providing professional development supports to teachers in both EHS and subsidized child care programs to ensure that teachers are interacting with children in ways that will support children's positive social-emotional development.

Contrary to hypotheses, no significant relationship was found between the EBS domain score and social-emotional competence on the BITSEA. It is not clear why there was no significant linear association; however, findings comport with prior studies using the CLASS-T with state samples of toddler classrooms that did not find significant relationships between the CLASS-T and social-emotional competence (La Paro et al., 2014). It could be that the current study sample, comprised of a mix of majority community-based child care and EHS programs, was more similar to these state samples, which were also comprised of more community-based programs that tend to be rated as

lower quality (Johnson et al., 2012; Loeb et al., 2004). If community-based programs tend to be lower quality, then the quality might not be high enough in this domain to be associated with children's positive social-emotional competence in the classrooms.

Future studies should examine the linear relationship between EBS and social-emotional competence in other samples of low-income toddler programs to identify whether there is a relationship between these two domains and to contribute to these mixed findings.

Thresholds in the Relationship Between the CLASS-T and Child Developmental Skills

As hypothesized, the findings suggested evidence of thresholds in the relationship between the CLASS-T and social-emotional problems and competence in this sample. The cut-points currently being used by Quality Counts aligned with thresholds in the a priori models examining the relationship between both CLASS-T domain scores and social-emotional skills, for both problems and social-emotional competence.

Findings for the models examining the relationship between both the EBS and ESL domain scores and social-emotional problems suggested that there was a threshold in both of these relationships. Results showed that there was a stronger negative relationship between the *EBS* domain score of the CLASS-T and *social-emotional problems* at higher levels of quality, above a quality cut-point of 5.50. Specifically, in classrooms that were rated at a 5.50 or higher on the EBS domain, there were lower classroom level social-emotional problems at higher levels of quality above a 5.50. However, in classrooms that were rated lower than a 5.50 on the EBS domain, there was no relationship between the EBS domain and social-emotional problems.

Alternatively, findings showed that in classrooms with lower levels of quality on the ESL domain (below a cut-point of 3.50), there was a stronger positive relationship between *ESL* and *social-emotional problems*. These findings suggest that within classrooms rated less than 3.50 on ESL, there were higher social-emotional problems in higher quality classrooms that were rated below a 3.50 on the ESL domain. Although these findings showed that there were more social-emotional problems seen at lower levels of quality, below a score of 3.50 on the ESL domain, this relationship was mitigated and no longer significant at higher levels of ESL, in classrooms rated as 3.50 or higher.

Findings for the models examining the relationship between both EBS and ESL and social-emotional competence also suggested that there were thresholds for both of these relationships. The results showed that there was a stronger positive relationship between both the *EBS* and *ESL* domains and *social-emotional competence* at higher levels of quality, above a quality cut-point of 5.50 for EBS and 3.50 for ESL. More specifically, within classrooms rated as 5.50 or higher on EBS and 3.50 or higher on ESL, there was more classroom level social-emotional competence at higher levels of quality. On the other hand, there was no relationship between either EBS or ESL and social-emotional competence at lower levels of quality.

Taken together, these findings provide initial empirical support for the cut-points currently being used in Quality Counts, as they align with thresholds in the relationship between the CLASS-T and social-emotional problems and competence where there are stronger relationships between the CLASS-T and social-emotional skills at different

levels of quality. Overall, these results support the idea that there is more positive social-emotional development at higher when compared to lower levels of quality.

Supplemental empirical analyses. To provide further empirical support for the a priori analyses, an empirical approach was applied by plotting the relationships between the CLASS-T domains and social-emotional problems and competence, and visually determining whether there was a nonlinear relationship and where the slope of the line changed. The results from the empirical approach showed that there were in fact nonlinear relationships between both CLASS-T domain scores and both BITSEA social-emotional problems and competence domains, based on the F test results. Additionally, when visually examining the regression line, the pattern indicated that there were inflection points where the slope of the line changed. These inflection points aligned with the a priori thresholds that were tested, providing an additional set of evidence to support the cut-points used in practice that align with thresholds in the relationship between the CLASS-T and social-emotional problems and competence.

The empirical approach supported and extended the findings of the a priori approach. The thresholds used in practice were visualized in the slopes for the empirical models. In addition, more than one threshold was identified in the relationship between ESL and both social-emotional problems and competence. Specifically, the results identified an additional threshold in the relationship between *ESL* and *social-emotional problems* that was not identified with the a priori approach where at higher levels of ESL, above 3.50, there was a negative relationship with social-emotional problems. This suggested that in classrooms rated above 3.50 on ESL, there were lower classroom level social-emotional problems. There was also an additional threshold in the relationship

between *ESL* and *social-emotional competence* that was not identified with the a priori approach, which showed that at lower levels of ESL, below 3.50, there was a negative relationship between ESL and social-emotional competence. These findings suggested that within classrooms rated lower than 3.50, there were lower classroom levels of social-emotional competence.

The findings from empirical models extend the a priori analysis results and indicate that there may be multiple thresholds in the relationship between ESL and children's social-emotional problems and competence. The additional thresholds may align with CLASS-T scores that are associated with other levels of Quality Counts ratings, such as at a level 5 or a level 3, for example. Future studies should extend these findings by using the a priori approach to identify if other thresholds exist at different cut-points on the CLASS-T that correspond to different rating levels of quality through Quality Counts, particularly for the relationship between ESL and social-emotional problems and competence.

Although the findings from the empirical approach both support and extend the a priori findings, it is difficult to determine exactly where the inflection points that indicate where the relationship between the CLASS-T domains and developmental skills changes and whether the a priori cut-points align perfectly with those inflection points. What is clear based on these graphs, however, is that children's developmental skills continue to improve as quality increases. Therefore, it is important to emphasize that meeting the pre-determined cut-points may not be enough and there should be a push for teachers to continue improving the quality of interactions in the classroom past the cut-points so that children can benefit from these interactions.

These findings support previous research conducted with the Baby FACES national sample of EHS programs that identified thresholds in the relationship between the CLASS-T and child developmental skills (Aikens et al., 2015). The cut-points tested in the current study were different from the cut-points used in the Baby FACES study; however, because this study examined a priori cut-points based on what was used in the local QRIS, Quality Counts. The Baby FACES study examined the cut-points empirically and identified various cut-points that aligned with thresholds, including cut-points that ranged between 4.5 and 5 on the EBS domain and between 3 and 4 on the ESL domain (depending on the child developmental skills examined in the study). Although the cut-points from both studies were similar, the purpose behind the analyses were different. The empirical approach in the Baby FACES study identified cut-points based on the data after it was analyzed and the a priori analyses in the current study verified already established thresholds based on cut-points used in practice.

The current study extends the findings from the national Baby FACES sample of EHS programs in several ways. First, both an a priori and empirical approach were used, to validate cut-points used in practice in the local QRIS program. The cut-points in the local program were comparable to the cut-points found in the national sample. However, the current study empirically tested the cut-points using the a priori approach so that findings were more relevant to inform what was actually being used in practice (Burchinal et al., 2016; Hatfield et al., 2016). Also, the current study examined the same cut-point across all children's developmental skills, rather than having the cut-points differ across developmental skills. Using multiple empirically identified cut-points may be more accurate statistically; however, it would not be practical in a systems level

program, such as QRIS, that relies on one overall cut-point score as a guideline for rating programs. Second, the current study sample included classrooms from EHS programs as well as classrooms from subsidized child care programs, extending the generalizability of prior findings.

Overall, findings support the idea that high quality teacher-child interactions characterized by both emotionally and behaviorally supportive interactions and those that support cognitive and language skills, may model positive relationships and support the development of behavior regulation, problem solving, and language skills, which contribute to more positive social-emotional skills in the classroom (La Paro et al., 2014; Mortensen & Barnett, 2015); and these associations are strongest at higher levels of quality.

Findings also provide evidence that the association between quality interactions and children's developmental skills may be non-linear and contributes to an important theoretical and statistical perspective in the field. Inconsistent linear associations between classroom quality and child skills are documented in several studies, and this may in part be because a non-linear association exists (Burchinal et al., 2011; Burchinal et al., 2016; Keys et al., 2013). Teacher-child interaction quality may need to reach a certain level before there is a strong, positive relationship between process quality and child developmental skills. These findings may also explain why there was no significant linear relationship found between EBS and social-emotional competence and a relationship at a trend level between ESL and social-emotional competence. The lack of a significant relationship between the CLASS-T domains and social-emotional competence could be the result of a nonlinear relationship between teacher-child interaction quality and social-

emotional competence in toddler samples and a non-linear approach should be considered in future studies examining this relationship.

Contrary to the study hypotheses, there was no evidence of either linear relationships or thresholds in the relationship between the CLASS-T and any of the developmental domains assessed by the ASQ-3, including language skills. These findings are in contrast to what was found with the Baby FACES sample, where a linear relationship was found between the ESL domain score and Communication subscale of the ASQ-3 (Bandel et al., 2014) and a non-linear relationship was found between the ESL domain score and the Problem Solving subscale of the ASQ-3 (Aikens et al., 2015).

There may be several methodological reasons why there were no significant relationships found between the CLASS-T and ASQ-3 scores. First, the ASQ-3 is a screener based on teacher-report of developmental domains. The ASQ-3 was used in the current study as it was the only available measure within the context of the larger studies to assess children's developmental skills. However, screening tools are limited to measure variation in children's developmental skills. Future studies should include other sources of measurement, such as direct assessments and observations, and comprehensive assessments, rather than screening tools, to assess children's developmental skills (Standards; American Educational Research Association [AERA], American Psychological Association [APA] & the National Council on Measurement in Education [NCME], 2014). Using measures that capture greater variation in children's skills across multiple domains, may make it easier to identify associations, both linear and non-linear, between teacher-child interaction quality and child developmental skills.

Second, the lack of relationships between teacher-child interaction quality and language skills might be because the CLASS-T is a global measure of teacher-child interaction quality. Several preschool studies have found stronger relationships between measures that assess more focused or specific interactions, such as the quality of interactions during language/literacy lessons, and children's language skills (Burchinal et al., 2016). Future studies could use measures to assess teacher-child interaction quality in specific contexts related to the developmental domain of interest, such as during language/literacy instruction, to identify whether interactions related to that domain may be more strongly associated with children's developmental skills.

Limitations and Future Directions

Although this study was one of the first to examine the validity of the CLASS-T scores and both the linear and nonlinear relationship between the CLASS-T and child developmental skills in both EHS and subsidized child care programs, there were some limitations. First, all of the child developmental skills were assessed using teacher-report. Although teachers can be reliable sources when reporting on children's developmental skills, their reports alone include error variance attributed to overall classroom differences, or teacher biases, rather than individual differences of children (Waterman, McDermott, Fantuzzo, & Gadsden, 2012). Future studies should assess child developmental skills using multiple methods of measurement and types of assessors, such as through direct assessment, observations, and parent-report, to obtain a more comprehensive assessment of children's developmental skills (Standards; AERA et al., 2014).

Second, the data collected and used for this study were collected concurrently, at one time point. The nature of the data collection limited the research design to one time point. Therefore, it was not possible to collect assessments at two different time points to examine change in developmental skills, and furthermore, the extent to which classroom process quality predicted changes in developmental skills. Additionally, there were no child or family background variables collected, such as parent education, employment or neighborhood risk factors; therefore, it was not possible to account for selection biases (e.g., any initial differences between children who selected into higher versus lower quality programs) in the models (Burchinal et al., 2016). Future studies could contribute to this work by controlling for selection factors and examining whether teacher-child interaction quality is related to change in developmental skills, above and beyond initial background characteristics. In addition, the CLASS-T was only observed at one time point, which gives us a snapshot of teacher-child interaction quality at one point in the year. However, research suggests that incorporating multiple time points of the CLASS-T would be important to examine as teacher-child interaction quality varies across the school year (Buell, Han, & Vukelich, 2017). In addition, future research could extend the current findings, by recording the activity setting and time of day during the CLASS observation, as preschool research suggests that there is also variability in the CLASS across a typical day and that this variability may be associated with children's developmental skills (Curby, Grimm, & Pianta, 2010).

A third limitation of this study was that variables that measured important proximal factors that directly influence young children's development, such as the family and home context, were not included. For toddlers, although high quality center-based

early educational environments are important, there are other fundamental proximal inputs, such as parent-child interactions and relationships, that have a strong influence on children's developmental skills at such an early age (McCartney et al., 2007). For example, one study found that the quality of the home environment predicted greater variance in toddlers' language and literacy skills, than the quality of the classroom (Pinto, Pessanha, & Aguiar, 2013). Additionally, other studies with preschool aged children in Head Start have shown that both classroom and family contexts uniquely and interactively predict child developmental skills (Bulotsky-Shearer, Wen, Faria, Hahs-Vaughn, & Korfmacher, 2012; Han, Schlieber, & Gregory, 2017; Wen, Bulotsky-Shearer, Hahs-Vaughn, & Korfmacher, 2012). Future studies could extend the current study findings by including variables that measure aspects of the home context, in order to examine the combined and interactive contribution of the home context and classroom quality that might account for variability in children's social-emotional and developmental skills in toddlers.

Finally, the current study sample size was limited to examine the measurement invariance of the CLASS-T across classrooms where teachers used predominantly English or Spanish. Research suggests that samples of at least 200 or more participants are needed to examine measurement invariance across different groups (Myers et al., 2011). Future studies could extend the initial findings by including a larger sample of linguistically diverse classrooms where teachers use Spanish in the classroom to identify whether the same published two-factor structure of the CLASS-T would be valid for use in linguistically diverse toddler classrooms.

Future research should also examine whether there is a bidirectional relationship between classroom quality and children's developmental skills and whether the composition of the classroom may influence classroom quality. This is important as research suggests that classroom composition influences the quality of teacher-child interactions (Buell et al., 2017). Additionally, research also shows that there is a bidirectional relationship between classroom quality and children's developmental skills, including child engagement in the classroom (Curby, Downer, & Booren, 2014). Identifying whether there is a bidirectional relationship between classroom quality and children's developmental skills, and whether classroom composition influences classroom quality in toddler classrooms, is particularly important in order to understand what factors may influence classroom quality and to consider these factors as they relate to teacher professional development around teacher-child interaction quality.

Implications for Policy and Practice

Findings from this study have implications for both policy and practice. Given increased federal and local policy support to invest in quality improvement initiatives (National Center on Early Head Start- Child Care Partnerships, 2016; Tout et al., 2009; Zaslow et al., 2010), it is imperative that the tools used to measure quality are validated for the purposes for which they are intended, as they are used to measure the quality of teacher-child interactions in the classroom to inform areas of quality improvement. Additionally, tools should be validated for use with the ethnically and linguistically diverse populations with whom they are being used. Findings from the current study extend the evidence for the validity of the CLASS-T scores when used to measure and monitor quality within programs such as the Miami-Dade EHS and Quality Counts

programs serving low-income, ethnically and linguistically diverse children. These findings lend support for the use of the scores to inform large scale national policy initiatives within EHS or QRIS that include teachers and children from diverse backgrounds.

The current study findings provide additional evidence that the CLASS-T scores relate in positive ways to toddler's social-emotional skills. This finding underscores the importance of high quality teacher-child interactions during the early childhood years, in particular as they support children's social-emotional development. One main concern in the broader early childhood field is the increase in social-emotional problems and challenging behaviors, particularly for children enrolled in under resourced child care centers (Division for Early Childhood of the Council for Exceptional Children, 2017). Recognizing the importance of resources for the professional development and support of early childhood teaching staff as part of these national policy initiatives are critical to improve classroom process quality, which in turn would support children's positive social-emotional development (Early, Maxwell, Ponder, & Pan, 2017; Finlon et al., 2015; Morris, Millenky, Raver, & Jones, 2013).

In addition, policy decisions are currently being made based on the CLASS-T cut-points that demarcate "low" versus "high" quality in diverse populations of teachers and children within community-based child care programs. Given the high stakes nature that is being applied to the use of the CLASS-T scores in this context, it is important to establish empirical evidence for the cut-points being used in practice to rate programs as either high or low quality. Validation of the cut-points, especially those used locally in quality improvement efforts in Florida, are especially important because they are linked

to eligibility of programs for funding and made public to parents. Considering the significant financial investments in quality improvement initiatives, such as Quality Counts, the current study findings add initial empirical evidence to support the use of cut-points in practice locally, although further studies should be examined using these cut-points before they are used for high-stakes purposes. Moreover, the current study findings also emphasize that meeting the pre-determined cut-points is not enough, as children's developmental skills continue to improve as classroom quality increases. These findings support the idea that it is not only important to meet the cut-points used in practice but imperative to exceed those cut-points so that children can benefit even more as the quality of teacher-child interactions increases. These findings also contribute to the idea that professional development for quality improvement in toddler classrooms serving low-income, diverse populations is especially important so that teachers are providing quality at a level that is beneficial for children's social-emotional development.

In summary, this study adds to the limited research findings that provide empirical support for the use of the CLASS-T domain scores to measure quality teacher-child interactions within programs serving low-income toddlers in EHS and subsidized child care programs. Study findings also contribute to the limited evidence base documenting the threshold at which CLASS-T domain scores more strongly relate to social-emotional development for toddlers. Findings underscore the need for greater resources to support early childhood teachers' knowledge of developmentally appropriate teaching practices and skills to engage in high quality emotionally supportive and engaging learning interactions within the classroom. As toddlerhood is one of the greatest times of rapid developmental change, when early skills are most malleable, setting a

target for reaching a level of quality is a critical opportunity for early childhood programs nationally to intervene so that the greatest long-term impact on young children is achieved.

References

- Aikens, N., Xue, Y., Bandel, E., Caronongan, P., Vogel, C. A., & Boller, K. (2015). *Early Head Start home visits and classrooms: Stability, predictors, and thresholds of quality*. Retrieved from https://www.acf.hhs.gov/sites/default/files/opre/early_head_start_home_visits_classrooms_stability_predictors.pdf.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA
- Bandel, E., Aikens, N., Vogel, C. A., Boller, K., & Murphy, L. (2014). *Observed quality and psychometric properties of the CLASS-T in the Early Head Start Family and Child Experiences Survey*. Retrieved from https://www.acf.hhs.gov/sites/default/files/opre/baby_faces_class_t_final_final_r.pdf.
- Bentler, P.M. (1990), "Comparative Fit Indexes in Structural Models," *Psychological Bulletin*, 107 (2), 238-46.
- Bodrova, E., Leong, D. J., Dickinson, D., & Neuman, S. (2006). Vygotskian perspectives on teaching and learning early literacy. *Handbook of Early Literacy Research*, 2, 243-256.
- Briggs-Gowan, M. J., & Carter, A. S. (2002). *Brief Infant-Toddler Social and Emotional Assessment (BITSEA) manual, version 2.0*. New Haven, CT: Yale University.
- Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology*, 29(2), 143-155.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 793-828). Hoboken, NJ, US: John Wiley & Sons Inc.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In: Bollen, K. A. & Long, J. S. (Eds.) *Testing Structural Equation Models*. pp. 136–162. Beverly Hills, CA: Sage.

- Brownell, C. A., & Kopp, C. B. (2007). Transitions in toddler socioemotional development. In C.A. Brownell & C.B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and transformations* (pp. 1-40). New York, NY: Guilford Publications.
- Buell, M., Han, M., & Vukelich, C. (2017). Factors affecting variance in Classroom Assessment Scoring System scores: season, context, and classroom composition. *Early Child Development and Care, 187*(11), 1635-1648.
- Bulotsky-Shearer, R. J., Wen, X., Faria, A. M., Hahs-Vaughn, D. L., & Korfmacher, J. (2012). National profiles of classroom quality and family involvement: A multilevel examination of proximal influences on Head Start children's school readiness. *Early Childhood Research Quarterly, 27*(4), 627-639.
- 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.
- Burchinal, M., Kainz, K., & Cai, Y. (2011). How well do our measures of quality predict child outcomes? A meta-analysis and coordinated analysis of data from large-scale studies of early childhood settings. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.) *Quality measurement in early childhood settings* (pp. 11-31). Baltimore, MD, US: Paul H Brookes Publishing.
- Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*(2), 166-176.
- Burchinal, M., Vernon-Feagans, L., Vitiello, V., Greenberg, M., & Investigators, F. L. P. K. (2014). Thresholds in the association between child care quality and child outcomes in rural preschool children. *Early Childhood Research Quarterly, 29*(1), 41-51.
- Burchinal, M., Zaslow, M., & Tarullo, L. (2016). *Quality thresholds, features, and dosage in early care and education: Secondary data analyses of child outcomes*. Wiley-Blackwell.
- Calkins, S. D. (2007). The emergence of self-regulation: Biological and behavioral control mechanisms supporting toddler competencies. In C.A. Brownell & C.B. Kopp (Eds.), *Socioemotional Development in the Toddler Years: Transitions and Transformations* (pp. 261-284). New York, NY: Guilford Publications.

- Calzada, E. J., Fernandez, Y., & Cortes, D. E. (2010). Incorporating the cultural value of respeto into a framework of Latino parenting. *Cultural Diversity and Ethnic Minority Psychology, 16*(1), 77.
- Carter, A.S., & Briggs-Gowan, M.J. (2006). *Manual for the infant-toddler social & emotional assessment (ITSEA) - Version 2*. San Antonio, TX: Psychological Corporation, Harcourt Press.
- Cassidy, D. J., Hestenes, L. L., Hansen, J. K., Hegde, A., Shim, J., & Hestenes, S. (2005). Revisiting the two faces of child care quality: Structure and process. *Early Education and Development, 16*(4), 505-520.
- Copple, C., & Bredekamp, S. (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*: ERIC.
- Curby, T. W., Downer, J. T., & Booren, L. M. (2014). Behavioral exchanges between teachers and children over the course of a typical preschool day: Testing bidirectional associations. *Early Childhood Research Quarterly, 29*(2), 193-204.
- Curby, T. W., Grimm, K. J., & Pianta, R. C. (2010). Stability and change in early childhood classroom interactions during the first two hours of a day. *Early Childhood Research Quarterly, 25*(3), 373-384.
- Dearing, E., McCartney, K., & Taylor, B. A. (2009). Does higher quality early child care promote low-income children's math and reading achievement in middle childhood? *Child Development, 80*(5), 1329-1349.
- Division for Early Childhood of the Council for Exceptional Children. (2017). *Position Statement on Challenging Behavior and Young Children*. Retrieved from https://challengingbehavior.cbcs.usf.edu/docs/DEC_PositionStatement_ChallengingBehavior.pdf.
- Early, D. M., Maxwell, K. L., Ponder, B. D., & Pan, Y. (2017). Improving teacher-child interactions: A randomized controlled trial of Making the Most of Classroom Interactions and My Teaching Partner professional development models. *Early Childhood Research Quarterly, 38*, 57-70.
- Feldman, R., & Klein, P. S. (2003). Toddlers' self-regulated compliance to mothers, caregivers, and fathers: implications for theories of socialization. *Developmental Psychology, 39*(4), 680.
- Finch, J. E., Johnson, A. D., & Phillips, D. A. (2015). Is sensitive caregiving in child care associated with children's effortful control skills? An exploration of linear and threshold effects. *Early Childhood Research Quarterly, 31*, 125-134.

- Finlon, K. J., Izard, C. E., Seidenfeld, A., Johnson, S. R., Cavadel, E. W., Ewing, E. K., & Morgan, J. K. (2015). Emotion-based preventive intervention: Effectively promoting emotion knowledge and adaptive behavior among at-risk preschoolers. *Development and Psychopathology*, 27(4 Pt 1), 1353–1365.
- Frey, W.H. (2018, March 14). *The US will become 'minority white' in 2045, Census projects* [Blog Post]. Retrieved from <https://www.brookings.edu/blog/the-avenue/2018/03/14/the-us-will-become-minority-white-in-2045-census-projects/>.
- Gonzalez-Mena, J., & Eyer, D. W. (2007). *Infants, toddlers, and caregivers: A curriculum of respectful, responsive care and education*. Boston, MA: McGraw-Hill.
- Gonzalez-Ramos, G., Zayas, L. H., & Cohen, E. V. (1998). Child-rearing values of low-income, urban Puerto Rican mothers of preschool children. *Professional Psychology: Research and Practice*, 29(4), 377.
- Gorsuch, R. L. (1983) *Factor analysis*. Hillsdale, NJ: Erlbaum.
- Greene, W. H. (2000). *Econometric analysis* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Guilamo-Ramos, V., Dittus, P., Jaccard, J., Johansson, M., Bouris, A., & Acosta, N. (2007). Parenting practices among Dominican and Puerto Rican mothers. *Social Work*, 52(1), 17-30.
- Hammer, C. S., Farkas, G., & Maczuga, S. (2010). The language and literacy development of Head Start children: A study using the Family and Child Experiences Survey database. *Language, Speech, and Hearing Services in Schools*, 41(1), 70-83.
- Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., & Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. *American Educational Research Journal*, 49(1), 88-123.
- Han, J., Schlieber, M., & Gregory, B. (2017). Associations of home and classroom environments with Head Start children's code-related and oral language skills. *Journal of Education for Students Placed at Risk (JESPAR)*, 22(4), 200-219.
- Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher-child interactions and preschool children's school readiness skills. *Early Childhood Research Quarterly*, 36, 561-571.

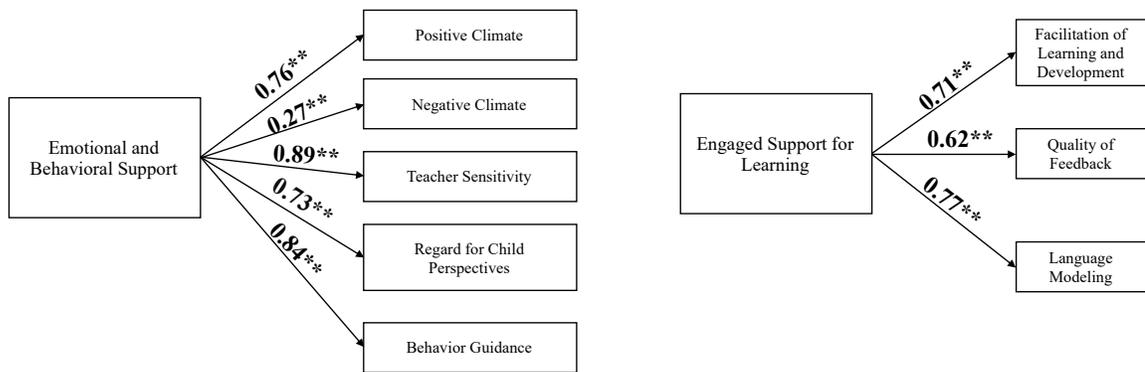
- Hatfield, B. E., Lower, J. K., Cassidy, D. J., & Faldowski, R. A. (2015). Inequities in access to quality early care and education: Associations with funding and community context. *Early Childhood Research Quarterly, 30*, 316-326.
- Hernandez, D., Denton, N., & McCartney, S. (2011). Early education programs: Differential access among young children in newcomer and native families *The next generation: Immigrant youth and families in comparative perspective*: Cornell University Press Ithaca, NY.
- Hestenes, L. L., Cassidy, D. J., Hegde, A. V., & Lower, J. K. (2007). Quality in inclusive and noninclusive infant and toddler classrooms. *Journal of Research in Childhood Education, 22*(1), 69-84.
- Hindman, A. H., Skibbe, L. E., Miller, A., & Zimmerman, M. (2010). Ecological contexts and early learning: Contributions of child, family, and classroom factors during Head Start, to literacy and mathematics growth through first grade. *Early Childhood Research Quarterly, 25*(2), 235-250.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117-144.
- 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.
- Hsu, N. (2017, February 20). *Highlights and Updates from QRIS* [Teachstone Blog]. Retrieved from <http://info.teachstone.com/blog/highlights-and-updates-from-qrisc>.
- Hu, L.T. and Bentler, P.M. (1999), Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6* (1), 1-55.
- Hungerford, G. M., Garcia, D., & Bagner, D. M. (2015). Psychometric evaluation of the brief infant-toddler social and emotional assessment (bitsea) in a predominately hispanic, low-income sample. *Journal of Psychopathology and Behavioral Assessment, 37*(3), 493-503.
- Jiang, Y., Granja, M., & Koball, H. (2017a). *Basic facts about low-income children, children under 3 years, 2015*. New York, NY: National Center for Children in Poverty, Mailman School of Public Health, Columbia University. Retrieved from: http://www.nccp.org/publications/pdf/text_1171.pdf.
- Jiang, Y., Granja, M. R., & Koball, H. (2017b). *Basic facts about low-income children: children under 6 years, 2015*. New York, NY: National Center for Children in Poverty, Mailman School of Public Health, Columbia University. Retrieved from http://www.nccp.org/publications/pdf/text_1172.pdf.

- Johnson, A. D., Ryan, R. M., & Brooks-Gunn, J. (2012). Child-care subsidies: Do they impact the quality of care children experience? *Child Development*, 83(4), 1444–1461.
- Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., . . . Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, 84(4), 1171-1190.
- Kisker, E. E., Boller, K., Nagatoshi, C., Sciarrino, C., Jethwani, V., Zavitsky, T., ... & Love, J. M. (2011). *Resources for measuring services and outcomes in Head Start programs serving infants and toddlers* (No. 742f4c265221408b8d1d70c965a92cc8). Mathematica Policy Research.
- La Paro, K. M., & Gloeckler, L. (2016). The context of child care for toddlers: The “Experience Expectable Environment”. *Early Childhood Education Journal*, 44(2), 147-153.
- La Paro, K., M, Hamre, B. K., & Pianta, R. C. (2012). *Classroom Assessment Scoring System: CLASS; Manual. Toddler*: Paul H. Brookes Publishing Company Baltimore, MD.
- La Paro, K. M., Williamson, A. C., & Hatfield, B. (2014). Assessing quality in toddler classrooms using the CLASS-Toddler and the ITERS-R. *Early Education and Development*, 25(6), 875-893.
- Layzer, J. I., & Goodson, B. D. (2006). The “quality” of early care and education settings: Definitional and measurement issues. *Evaluation Review*, 30(5), 556-576.
- Le, V.-N., Schaack, D. D., & Setodji, C. M. (2015). Identifying baseline and ceiling thresholds within the Qualistar early learning quality rating and improvement system. *Early Childhood Research Quarterly*, 30, 215-226.
- Li, W., Farkas, G., Duncan, G. J., Burchinal, M. R., & Vandell, D. L. (2013). Timing of high-quality child care and cognitive, language, and preacademic development. *Developmental Psychology*, 49(8), 1440.
- Loeb, S., Fuller, B., Kagan, S. L., & Carrol, B. (2004). Child care in poor communities: Early learning effects of type, quality, and stability. *Child Development*, 75(1), 47-65.
- Magnuson, K. A., & Duncan, G. J. (2006). The role of family socioeconomic resources in the black–white test score gap among young children. *Developmental Review*, 26(4), 365-399.
- Marsh, L. C., & Cormier, D. (2002). *Spline regression models*. London, UK: Sage.

- 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.
- McCartney, K., Dearing, E., Taylor, B. A., & Bub, K. L. (2007). Quality child care supports the achievement of low-income children: Direct and indirect pathways through caregiving and the home environment. *Journal of Applied Developmental Psychology, 28*(5), 411-426.
- Morris, P., Millenky, M., Raver, C. C., & Jones, S. M. (2013). Does a preschool social and emotional learning intervention pay off for classroom instruction and children's behavior and academic skills? Evidence from the foundations of learning project. *Early Education and Development, 24*(7), 1020-1042.
- Mortensen, J. A., & Barnett, M. A. (2015). Teacher-child interactions in infant/toddler child care and socioemotional development. *Early Education and Development, 26*(2), 209-229.
- Muthén, L. K., & Muthén, B. O. (1998-2011). *MPlus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Myers, N. D., Ahn, S., & Jin, Y. (2011). Sample size and power estimates for a confirmatory factor analytic model in exercise and sport: A Monte Carlo approach. *Research Quarterly for Exercise and Sport, 82*(3), 412-423.
- National Center on Early Head Start-Child Care Partnerships. (2016). *Growing the supply of Early Learning Opportunities for Infants and Toddlers*. Retrieved from <http://www.acf.hhs.gov/programs/eecd/early-learning/ehs-cc-partnerships>.
- National Scientific Council on the Developing Child (2010). *Early experiences can alter gene expression and affect long-term development: Working paper No. 10*. Retrieved from www.developingchild.harvard.edu.
- NICHD, E. (2002). Early child care and children's development prior to school entry: Results from the NICHD study of early child care. *American Educational Research Journal, 39*(1), 133-164.
- Office of Child Care. (2014). *Reauthorization of the Child Care and Development Fund (CCDF): An exciting new era for child care*. Retrieved from http://www.acf.hhs.gov/programs/occ/ccdf_reauthorization.
- Piaget, J. (1951). *The child's conception of the world*: Rowman & Littlefield.
- Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems. *The Future of Children, 26*(2), 119-137.

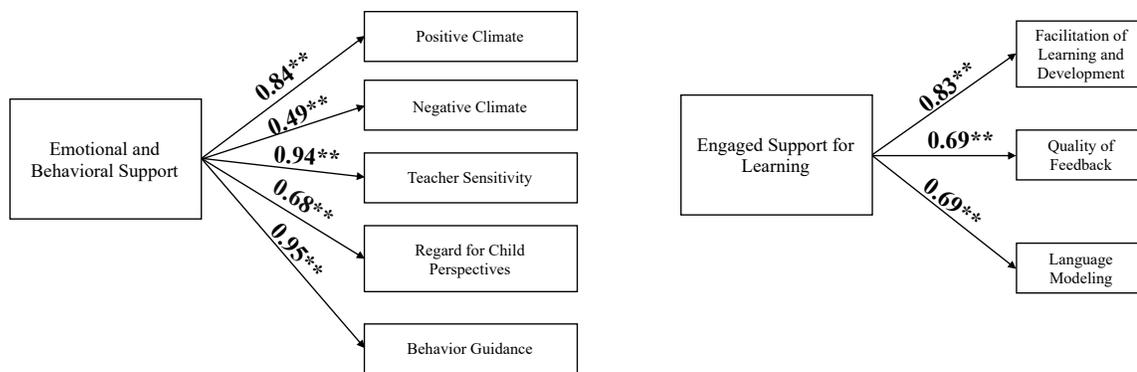
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science, 9*(3), 144-159.
- Pianta, R. C., La paró, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS) manual, pre-K*: Paul H. Brookes Publishing Company Baltimore, MD.
- Pianta, R.C., La Paro, K.M., & Hamre, B.K. (2010) *Classroom Assessment Scoring System Pilot Toddler Manual (CLASS-T)*. Charlottesville, VA: Teachstone, Inc.
- Pinto, A. I., Pessanha, M., & Aguiar, C. (2013). Effects of home environment and center-based child care quality on children's language, communication, and literacy outcomes. *Early Childhood Research Quarterly, 28*(1), 94-101.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Ruzek, E., Burchinal, M., Farkas, G., & Duncan, G. J. (2014). The quality of toddler child care and cognitive skills at 24 months: Propensity score analysis results from the ECLS-B. *Early Childhood Research Quarterly, 29*(1), 12-21.
- Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher-child relationships. *Attachment & Human Development, 14*(3), 213-231.
- Schaack, D., Tarrant, K., Boller, K., & Tout, K. (2012). Quality rating and improvement systems: Frameworks for early care and education systems change. In S.L. Kagan & K. Kauerz (Eds.), *Early Childhood Systems: Transforming Early Learning* (pp. 71-86). New York, NY: Teachers College Press.
- Setodji, C. M., Le, V.-N., & Schaack, D. (2013). Using generalized additive modeling to empirically identify thresholds within the ITERS in relation to toddlers' cognitive development. *Developmental Psychology, 49*(4), 632.
- Squires, J., Twombly, E., Bricker, D., & Potter, L. (2009). *ASQ-3: Users Guide*. Baltimore: Paul H.
- Steiger, J. H. (1990). Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioural Research, 25*, 173 - 180.

- Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the next decade of quality rating and improvement systems*. Washington, DC: Child Trends.
- U.S. Census Bureau. (2013). *Who's minding the kids? Child care arrangements: Spring 2011*. Retrieved from <http://www.census.gov/prod/2013pubs/p70-135.pdf>.
- Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development, 81*(3), 737-756.
- Vesely, C. K. (2013). Low-income African and Latina immigrant mothers' selection of early childhood care and education (ECCE): Considering the complexity of cultural and structural influences. *Early Childhood Research Quarterly, 28*(3), 470-486.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*: Cambridge, MA: Harvard University Press.
- Watanabe, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double jeopardy: Poorer social-emotional outcomes for children in the NICHD SECCYD experiencing home and child-care environments that confer risk. *Child Development, 82*(1), 48-65.
- Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2012). The matter of assessor variance in early childhood education—Or whose score is it anyway?. *Early Childhood Research Quarterly, 27*(1), 46-54.
- Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. *Early Childhood Research Quarterly, 28*(2), 199-209.
- Wen, X., Bulotsky-Shearer, R. J., Hahs-Vaughn, D. L., & Korfmacher, J. (2012). Head Start program quality: Examination of classroom quality and parent involvement in predicting children's vocabulary, literacy, and mathematics achievement trajectories. *Early Childhood Research Quarterly, 27*(4), 640-653.
- Whitebook, M., McLean, C., Austin, L.J.E., & Edwards, B. (2018). *Early Childhood Workforce Index – 2018*. Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley. Retrieved from <http://csce.berkeley.edu/topic/early-childhood-work-force-index/2018/>.
- Zaslow, M., Tout, K., & Martinez-Beck, I. (2010). *Measuring the quality of early care and education programs at the intersection of research, policy, and practice*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, and US Department of Health and Human Services.



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

Figure 1. Factor Loadings from Final Two-Factor Model of CLASS-T with Overall Sample



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

Figure 2. Factor Loadings from CFA of Two-Factor Model of CLASS-T in Classrooms that Use Predominantly English

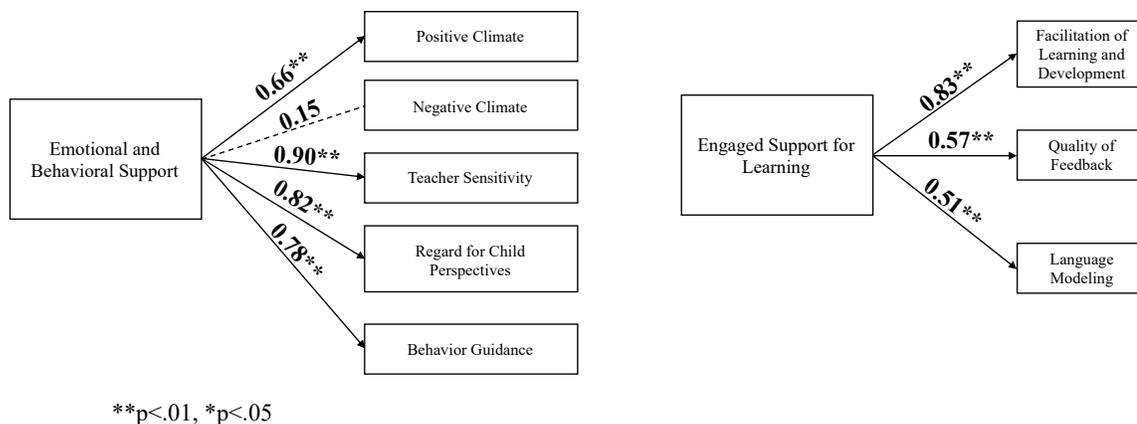


Figure 3. Factor Loadings from CFA of Two-Factor Model of CLASS-T in Classrooms that Use Predominantly Spanish

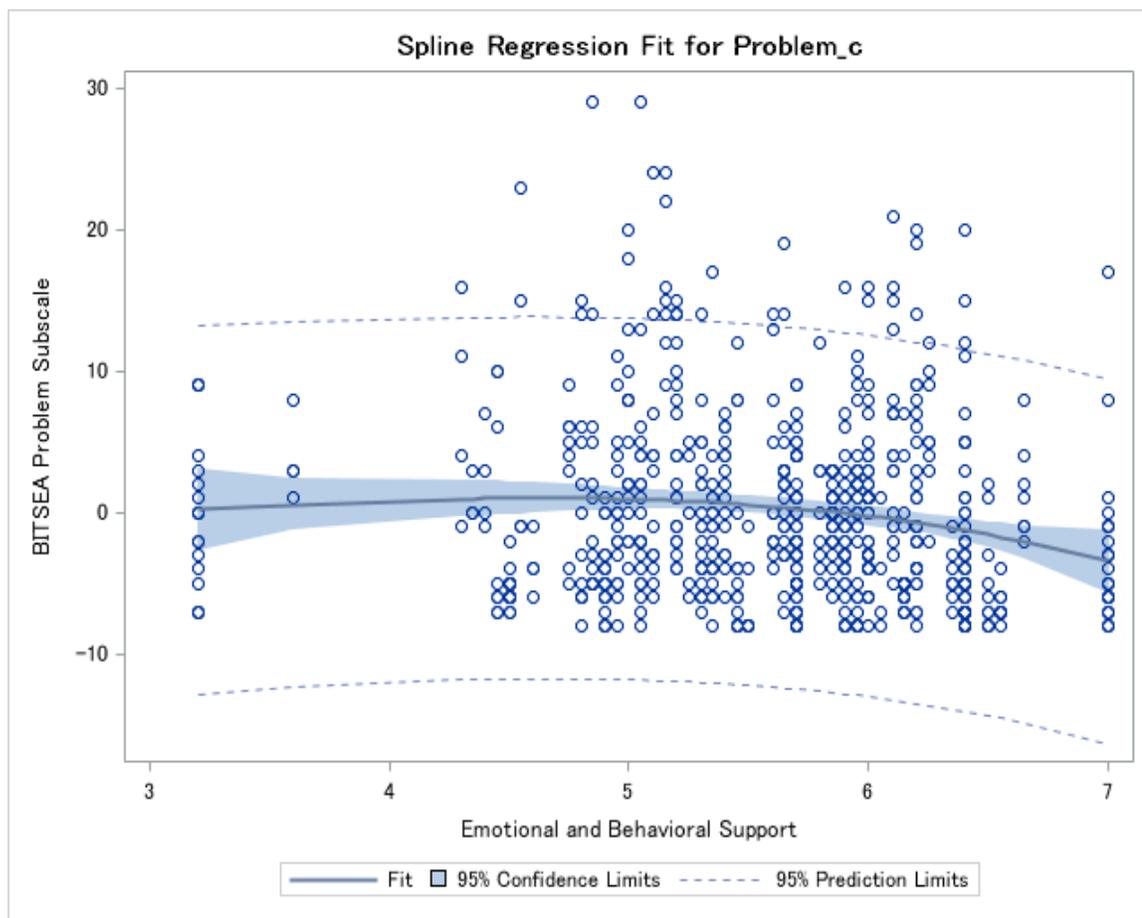


Figure 4. Nonlinear Relationship between EBS and BITSEA Social-emotional Problems
Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment.

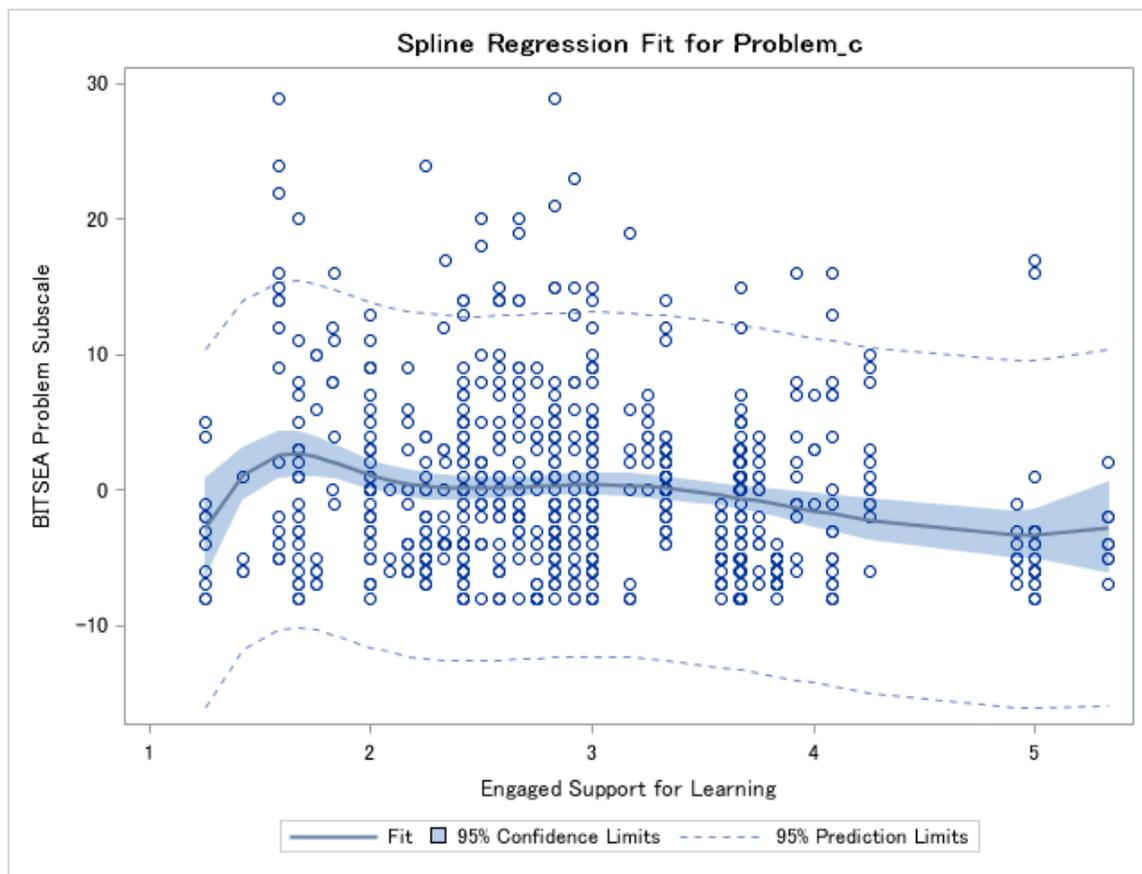


Figure 5. Nonlinear Relationship between ESL and BITSEA Social-emotional Problems
Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment.



Figure 6. Nonlinear Relationship between EBS and BITSEA Social-emotional Competence

Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment.

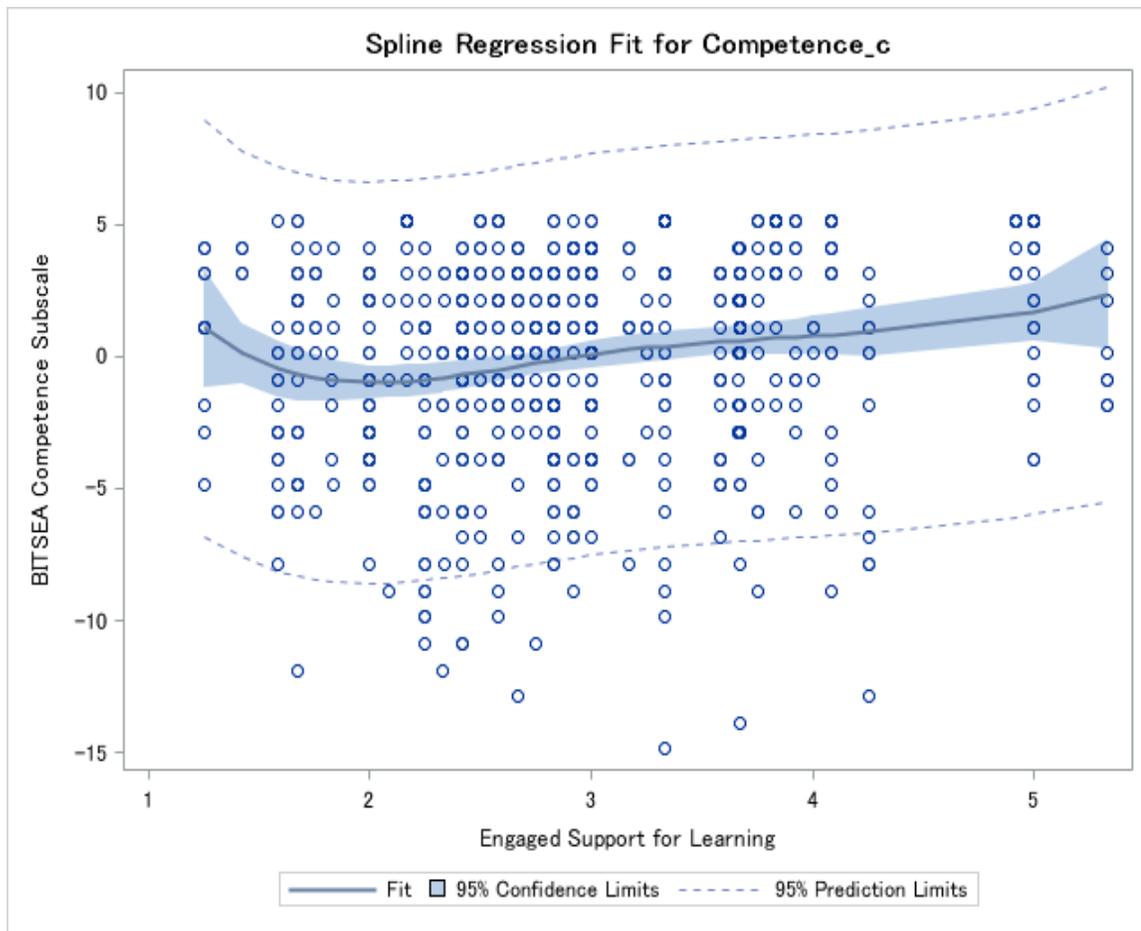


Figure 7. Nonlinear Relationship between ESL and BITSEA Social-emotional Competence

Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment.

Table 1
Means and Standard Deviations for Measures

	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>Range</i>
<i>CLASS-T</i>				
Emotional and Behavioral Support (EBS)	106	5.51	0.69	3.20-7.00
Engaged Support for Learning (ESL)	106	2.84	0.84	1.17-5.33
<i>BITSEA</i>				
Social-emotional Problems	669	8.02	6.57	0-37
Social-emotional Competence	669	16.87	3.92	2-22
<i>ASQ-3</i>				
Communication	262	48.07	15.62	0-60
Gross Motor	262	55.88	7.59	0-60
Fine Motor	262	49.26	12.10	0-60
Problem Solving	261	49.87	11.69	0-60
Personal-social	261	49.63	11.43	0-60

Note. CLASS-T = Classroom Assessment Scoring System-Toddler, BITSEA= Brief Infant-Toddler Social and Emotional Assessment, ASQ-3= Ages and Stages Questionnaires.

Table 2
Bivariate Correlations between Measures

	1	2	3	4	5	6	7	8	9
1. CLASS-T EBS	—	.57**	-.12**	.08*	-.05	-.10	-.02	.00	-.02
2. CLASS-T ESL		—	-.15**	.16**	.02	-.13*	-.03	.04	-.01
3. BITSEA Problems			—	-.40**	-.18**	-.01	-.10	-.10	-.14*
4. BITSEA Competence				—	.40**	.28**	.32**	.32**	.26**
5. ASQ-3 Communication					—	.41**	.41**	.64**	.54**
6. ASQ-3 Gross Motor						—	.46**	.51**	.45**
7. ASQ-3 Fine Motor							—	.64**	.44**
8. ASQ-3 Problem Solving								—	.62**
9. ASQ-3 Personal-social									—

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

Note. CLASS-T = Classroom Assessment Scoring System-Toddler, BITSEA= Brief Infant-Toddler Social and Emotional Assessment, ASQ-3= Ages and Stages Questionnaires.

Table 3
Linear Relationships between CLASS-T and BITSEA Domains

	<i>BITSEA Problems</i>	<i>BITSEA Competence</i>
	<i>B (SE)</i>	<i>B (SE)</i>
Intercept	14.64** (3.84)	13.62** (2.24)
Age	0.02 (0.04)	0.11** (0.03)
Sex	1.47** (0.42)	-1.02** (0.25)
Emotional and Behavioral Support	-1.40* (0.66)	0.25 (0.38)
Engaged Support for Learning	-1.24* (0.54)	0.60 (0.31)

N=667. ** $p < .01$, * $p < .05$.

Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment.

Table 4
 Linear Relationships between CLASS-T and ASQ-3 Domains

	ASQ-3	ASQ-3	ASQ-3	ASQ-3	ASQ-3	ASQ-3
	Communication	Gross Motor	Fine Motor	Problem Solving	Personal-social	
	B (SE)	B (SE)				
Intercept	39.77** (9.67)	57.71** (4.00)	61.09** (7.66)	42.67** (7.09)	40.74** (8.16)	
Age	0.42** (0.14)	0.09 (0.06)	-0.25* (0.11)	0.19 (0.10)	0.22* (0.11)	
Sex	-7.11** (1.80)	-1.09 (0.94)	-4.97** (1.42)	-3.45* (1.41)	-3.90** (1.28)	
Emotional and Behavioral Support	0.19 (1.58)	-0.65 (0.63)	-0.65 (1.25)	0.76 (1.15)	0.92 (1.35)	
Engaged Support for Learning	1.61 (1.60)	-1.21 (0.71)	-1.18 (1.28)	1.42 (1.19)	0.78 (1.35)	

N=260. ** $p < .01$, * $p < .05$.

Note. ASQ-3= Ages and Stages Questionnaires.

Table 5
Testing for Quality Thresholds using A Priori Cut-points of CLASS-T

Cut-points	Emotional and Behavioral Support			Engaged Support for Learning		
	Low	High	differ?	Low	High	differ?
	<i>B (SE)</i>	<i>B (SE)</i>		<i>B (SE)</i>	<i>B (SE)</i>	
BITSEA Problems	0.02 (0.15)	-2.56** (0.90)	L<H**	0.65* (0.27)	-1.09 (0.86)	L>H*
BITSEA Competence	0.15 (0.09)	2.00** (0.54)	L<H**	-0.03 (0.16)	1.46** (0.51)	L<H**
ASQ-3 Communication	-1.16* (0.58)	-5.95 (4.38)	ns	-0.01 (1.13)	1.04 (4.40)	ns
ASQ-3 Gross Motor	-0.50 (0.29)	-4.47* (2.22)	ns	1.42* (0.57)	1.36 (2.20)	ns
ASQ-3 Fine Motor	-0.25 (0.46)	-2.08 (3.49)	ns	0.85 (0.89)	1.35 (3.47)	ns
ASQ-3 Problem Solving	0.03 (0.45)	2.99 (3.40)	ns	0.71 (0.87)	5.47 (3.38)	ns
ASQ-3 Personal-social	-1.24** (0.43)	-6.37 (3.26)	ns	1.70* (0.84)	2.40 (3.27)	ns

BITSEA, N=667; ASQ-3, N= 260. ** $p < .01$. * $p < .05$.

Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment, ASQ-3= Ages and Stages Questionnaires.

Table 6
Testing for Quality Thresholds Empirically using B-spline Analyses: Nonlinear F Tests

	<i>Emotional and Behavioral Support</i>	<i>Engaged Support for Learning</i>
	<i>Nonlinear F</i>	<i>Nonlinear F</i>
BITSEA Problems	6.23**	5.44**
BITSEA Competence	7.99**	8.42**

N=667. ** $p < .01$, * $p < .05$.

Note. BITSEA= Brief Infant-Toddler Social and Emotional Assessment.