

An Equity Perspective on Measurement Invariance and Profiles of Head Start Classroom  
Quality: Implications for School Readiness, Policy, and Practice

by

Brittany Alexander

A Dissertation Presented in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

Approved June 2022 by the  
Graduate Supervisory Committee:

Hyung Chol (Brandon) Yoo, Co-Chair  
Shantel Meek, Co-Chair  
Kelly Edyburn  
Manuela Jimenez Herrera

ARIZONA STATE UNIVERSITY

August 2022

## ABSTRACT

Quality in early childhood education (ECE) is central to equitable child development and preparation for formal schooling and has been widely studied by researchers and of interest to policy makers. As the federal pre-k program, Head Start is a key ECE context to understand quality and its implications for equity. One central measure of classroom quality, the Classroom Assessment Scoring System (CLASS), is used in policy-making and funding decisions to study the impact of quality on children's school readiness. The CLASS is a measure of teacher-child interactional quality, but measurement invariance across teacher race/ethnicity has yet to be examined for this measure in the published literature. Additionally, patterns of classroom quality and the sociocultural context of classrooms as predictors of children's social skills and approaches to learning have yet to be examined. Using anti-racist early childhood education theory and a nationally representative Head Start sample, the Family and Child Experiences Survey 2009 cohort, I conducted two studies to address these gaps. In the first study, I investigated the measurement invariance of the CLASS across teacher race/ethnicity (Black, Latine, White). I found evidence of partial strong invariance, with only one non invariant parameter for Black teachers, suggesting that means may be compared across teacher race/ethnicity. However, the implications of these findings must be interpreted through an equity lens, and quality measures should work to include equity indicators explicitly. In the second study, I examined patterns of classroom quality indicated by the CLASS and 1) dual language learner (DLL) composition and 2) in combination with child demographics and teacher-child demographic match as predictors of school readiness outcomes. I found evidence of three profiles of classroom quality and

DLL composition did not significantly predict profile membership. Further the profile with higher levels of negative climate and moderate emotional support and classroom organization negatively predicted child social skills and approaches to learning. Applying anti-racist ECE theory studies suggest that the CLASS does not sufficiently address equity in ECE, but may be used with Black, Latine, and White teachers and low quality should be addressed through intervention to prevent negative outcomes for children.

## ACKNOWLEDGMENTS

This dissertation would not have been possible without my support system, my mentors, colleagues, friends, family, and partner. Collectively this group has provided me with a compassionate space to explore ideas, be vulnerable, and pushed me to learn and grow. In particular I would like to thank:

### Mentors

Dr. Shantel Meek, for sharing your experiences and expertise in ECE policy, for believing in me, for always challenging me to think in actionable and solutions-oriented ways, and for giving me the opportunity to pursue meaningful work with wonderful colleagues post-graduation.

Dr. Brandon Yoo, your mentorship allowed me to finish my doctorate, to work out my long-term goal of applying critical theory to teacher-child interaction focused content. Thank you for taking a chance on me and mentoring me late in my program, including me in your lab. Your support allowed me to finish my degree.

Dr. Kelly Edyburn, for being incredibly thoughtful and generous with your time and mentorship. Your critical approach to research and early childhood work has inspired me in so many ways.

Dr. Manuela Jimenez Herrera, for your patience and kindness over the years as I grew and changed throughout graduate school. Thank you for always pushing me to be better.

Dr. Roopa Iyer, for the opportunity to learn and apply my work in the context of Arizona ECE policy. You allowed me to pursue such meaningful work and welcomed me into an incredible evaluation team.

Dr. Casey Sechler, for your friendship, mentorship, and continuous support. Thank you for your patience and thoughtful mentorship throughout the last 6 years.

Drs. Crystal Bryce, Tashia Abry, for giving me the space to make mistakes and develop the skills I needed pursue a career in research.

Dr. Richard Fabes, for taking a chance on me as I entered graduate school and supporting my varied interests and graduate school journey.

Drs. Zachary Casey and Marsha Walton, for their transformative teaching, research, and mentorship that inspired me to pursue my doctorate.

#### Friends and Colleagues

To my graduate school friends and colleagues, Dr. Diana Jenkins, Jayley Janssen, Dr. Annabelle Atkin, Dr. Michelle Pasco, Claire Abidog, thank you for your endless support, compassion, and laughter. I truly treasure our friendships and would not have made it through this process without you.

To my lab mates Abigail Gabriel, Abby Saavedra, Ronae Matriano, and Sarah Parks, thank you for welcoming me into your academic space and for sharing your work with me. You all inspire me.

#### Family

My parents Amy and Troy Alexander, thank you for always believing in me and supporting my hopes and dreams.

To my siblings (Heidi, T.J., and Tucker), my grandparents (Gayle and John Edgar, Annie and Bob Dallow, and Wayne and Laura Jane Alexander), and my in-laws (Paddy and Sam Wong, Liz and Shane Roland, and Matthew Wong and Josie Benson), thank you all for your love and support.

To my husband and partner, Daniel Wong. Your love, patience, support, and character have made my graduation possible. I am endlessly grateful to you.

Finally, I dedicate this dissertation to my mother, who attended my defense from her hospital bed, and passed less than two weeks later. Your compassion, love, and commitment will stay with me always.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER	
1 GENERAL INTRODUCTION .....	1
Importance of Head Start .....	2
Definitions .....	6
Theoretical Perspectives.....	7
Present Studies Overview .....	10
2 STUDY 1 INTRODUCTION.....	13
Present Study .....	26
3 STUDY 1 METHOD .....	27
Participants .....	27
Measures .....	27
4 STUDY 1 RESULTS .....	31
Preliminary Analyses .....	31
Primary Analyses .....	31
5 STUDY 1 DISCUSSION .....	35
Summary of Findings .....	35
Limitations and Future Directions .....	41
6 STUDY 2 INTRODUCTION.....	43

CHAPTER	Page
Present Study .....	57
7 STUDY 2 METHOD .....	59
Participants .....	59
Measures .....	60
8 STUDY 2 RESULTS .....	64
Preliminary Analyses .....	64
Primary Analyses .....	64
9 STUDY 2 DISCUSSION .....	71
10 GENERAL DISCUSSION .....	82
REFERENCES .....	88
APPENDIX	
A STUDY 1 & 2 TABLES .....	102
B STUDY 2 FIGURES .....	113



LIST OF TABLES

Table	Page
1. Study 1 Correlations Among Variables .....	102
2. Study 1 Descriptive Statistics.....	103
3. Teacher Invariance Global Fit Indices.....	104
4. CLASS Standardized Factor Loadings Across Measurement Invariance .....	105
5. Correlations Among Level 2 Variables .....	106
6. Study 2 Descriptive Statistics .....	107
7. Sample Teacher Child Racial/Ethnic Match .....	107
8. Classification Criteria for Classroom Quality Profiles .....	108
9. Quality Profiles Across Sample Statistics .....	109
10. Sample Race/Ethnicity Within Quality Profile .....	110
11. Quality and Context Predictors of Social Skills and Approaches to Learning.	111

## LIST OF FIGURES

Figure		Page
1.	Quality Profiles 2-Class Solution (Mean-Centered) .....	113
2.	Quality Profiles 3-Class Solution (Mean-Centered).....	114
3.	Quality Profiles 4-Class Solution (Mean-Centered).....	114
4.	Quality Profiles 3-Class Solution.....	115

# CHAPTER 1

## GENERAL INTRODUCTION

The early years are the foundation of children's futures and are central to creating equitable opportunities for children and families to thrive (Britto et al., 2017). Early childhood education (ECE) can help to facilitate the development of school readiness, positive peer relationships, and positive teacher-child relationships, which can help to set children up well for school and life (Lifter et al., 2011; Yoshikawa et al., 2013). Head Start, the federal pre-kindergarten (pre-K) program, was a victory of activists in the Civil Rights movement and continues to serve working-class children and families. It is also structured as a two-generation program that focuses its wrap-around services on families in addition to children, providing job opportunities and training to Head Start parents. During its early years Civil Rights activists harnessed Head Start to help fight segregation and structural racism (Sanders, 2016). For example, in places like Mississippi, employment and educational opportunities were frequently denied to Black Mississippians. However, and as a federally funded program, Head Start enforced nondiscrimination laws. Many Black Mississippians, largely Black women, harnessed this opportunity for employment and influence (Sanders, 2016). Additionally, the state curriculum at the time glorified the Confederacy and white supremacy, but in Head Start settings Black educators were able to teach Black history and support holistic education of Black children (Sanders, 2016).

There have always been advocates for antiracism in education, however, the history and pervasiveness of racism in the United States continues to evolve and curtails

children’s opportunities to receive an equitable education before they even enter the classroom (Ladson-Billings, 2013). Nationally, Head Start provides access to ECE for approximately 873,000 children, yet there are racial disparities in access to quality education, including Head Start centers (Hardy et al., 2020; Head Start Early Childhood Learning & Knowledge Center, 2021). For example, a study in Georgia found that communities with higher Black and Hispanic populations tended to have low rated pre-K programs (Bassok & Galdo, 2016). Nationally, Black and Hispanic children tend to live in neighborhoods with far more Head Start eligible children per neighborhood center than White children (On average 90 eligible Black children, 100 eligible Hispanic children, and 60 eligible White children; Hardy et al., 2020). This means that Black and Hispanic children are more likely than White children to live in neighborhoods where the demand for Head Start outweighs the supply, underscoring inequitable access to Head Start centers for Black and Hispanic children (Hardy et al., 2020). Nationally only 38% of Hispanic and 54% of Black children that are eligible for Head Start are enrolled (Hardy et al., 2020). However, little is known about disparities in access to high-*quality* ECE (i.e., the quality of interactions between children and teachers as defined by the Classroom Assessment Scoring System [CLASS]) within Head Start or the potential bias of high-stakes measures used to assess quality in the ECE workforce.

### **Importance of Head Start, Quality, and School Readiness**

In the following sections I provide an overview of the research on Head Start, quality, and school readiness and discuss their policy importance and context.<sup>1</sup> School

---

<sup>1</sup> Researcher Positionality Statement: My positionality as a White able-bodied woman, raised in a protestant, upper middle class context informs my experiences and therefore the lens through which I view

readiness refers to the skills that children need to be able to function successfully in the formal schooling context. It has been shown to predict future success in school, including mental health, school adjustment, and academic achievement (Eisenberg et al., 2005; Quirk et al., 2013; Ricciardi et al., 2021). The early development of these skills has been shown to have long-term impacts. For example, Ricciardi and colleagues (2021) found that children's school readiness at age four predicted children's academic achievement through the entirety of elementary school. Quality ECE is thought to be one such mechanism through which school readiness skills are developed.

## **Research**

Robust research findings suggest that for early childhood education to be beneficial for children's school readiness outcomes, it must be high-quality (Burchinal et al., 2016; Zaslow et al., 2010). Quality ECE has been shown to promote children's healthy development across nearly every developmental domain including cognitive, social, emotional, and physical (McClelland et al., 2017; Vandenbroucke et al., 2018; Venetsanou & Kambas, 2010; Yoshikawa, 2013). Head Start is widely considered to be a high quality ECE provider, in part due to the high standards grantees are held to at the federal level to maintain funding. It is also the largest ECE program in the U.S., serving roughly 873,000 children in 2019 (Head Start Early Childhood Learning & Knowledge Center, 2021). Therefore, researchers have extensively examined the impact of Head Start on children, families, and communities (Lee et al., 2021; Puma et al., 2010). However, only an estimated 50% of eligible children have access to Head Start (Schmit

---

the world. Though I attempt to actively interrogate my biases and approach my research through a lens of cultural humility, my social location is inextricably linked to my research.

& Walker, 2016). Even for those children who obtain access to Head Start, there is variability in quality within Head Start (Morris et al., 2018; Puma et al., 2010).

Head Start takes a federal-to-local approach, provides wrap-around services for families (i.e., job training, health insurance, developmental screening), and offers child care coverage for working class families and those in poverty. Researchers have largely found positive associations between Head Start participation and children's outcomes, including larger gains for dual language learners (DLLs; Bierman et al., 2008a; Bierman et al., 2008b; Lee, 2019; Lee et al., 2021). Studies have primarily examined children's social, emotional, and cognitive outcomes (Lee et al., 2021). The effects of Head Start ECE are stronger, proximally closer to participation (Pages et al., 2020). This may be due to treatment effects and or the quality of the educational contexts that children enter once they leave Head Start (Pages et al., 2020). Better understanding of the way that quality manifests in Head Start, particularly patterns (i.e., not simply examining overall quality, but the potential differential impact of various levels of quality indicators). Examining patterns of quality within Head Start settings and their relation to child outcomes may provide relevant information about the key ingredients of Head Start that help to improve children's school readiness.

## **Policy**

There are many federal and state-level policy efforts to address children's access to quality ECE, including Head Start, the Child Care Development Block Grant, the Preschool Development Grant (B-5), Early Head Start Child Care Partnerships, Race to the Top - Early Learning Challenge, and quality rating and improvement systems (QRIS).

Policy considerations for allocating ECE funds and implementing these policy efforts must balance competing priorities. For example, funding decisions must balance expanding access to ECE, improving quality of ECE, and ideally expanding access to quality ECE. Further, the importance of observing demonstrated improvements in outcomes as a result of these investments looms large. Tools such as the CLASS are used to quantify the quality of the investment of public dollars into ECE context. At the federal level, this means that the CLASS is used to assess Head Start quality in an attempt to understand the context of child outcomes.

Quality within Head Start and in ECE settings across the country is commonly measured by the CLASS. This measure does not capture equity but is a central measure of classroom quality used to hold Head Start grantees to a high standard of quality and to ensure the continuation of their funds. The high-stakes nature of the measure for funding decisions also impacts the Head Start workforce, which is widely overworked and underpaid, therefore it is important that the measure is not biased on demographic characteristics such as race/ethnicity (Delaney & Krepps, 2021; McLean et al., 2019). According to a review of the literature and the existing validity evidence of the measure, measurement invariance of the CLASS across teacher race/ethnicity has yet to be tested (McDoniel et al., 2022; Perlman, 2016). Additionally, patterns of classroom quality, the sociocultural context of classrooms, and the implications for school readiness gaps in the literature that remain to be investigated. Therefore, I propose two studies 1) to investigate measurement invariance across teacher race/ethnicity on the CLASS and 2) to empirically investigate patterns of quality in Head Start, examine the demographic composition of these profiles, DLL composition as a predictor of profile membership,

and predict children's school readiness. First, I provide key definitions and theoretical perspectives guiding this proposal before detailing the two proposed studies.

## **Definitions**

There are four key terms to define for this proposal. First, the overarching structure of racism, or *white supremacy*, refers to a system of racial hierarchy that serves to preserve white domination through the marginalization of people of color (DiAngelo, 2018). This definition is closely tied to the definition of *racism*, or a system of racial hierarchy enacted through the privileging of some groups and oppression of others, which operates at the structural, institutional, organizational, and interpersonal level (Golash-Boza, 2016; Jones, 1997; Juang et al., 2017). At all levels, racism serves to benefit white supremacy. Oftentimes the definition of racism perpetuated particularly by White people is limited to unfair treatment of one individual by another based on racial prejudice and the definition of white supremacy is tied to extremist groups such as the KKK and the Proud Boys (DiAngelo, 2018). However, these definitions do not adequately capture the pervasiveness of racism and serve to limit racism to the bad actions of a few individuals rather than a system of oppression that operates across multiple levels. However, the opposite of racism is not simply to be not racist, as when we understand racism as a multi-level power structure, being not racist is simply beyond individual control. Rather, being explicitly and actively against these racist structures, or *anti-racist*, offers a way to conceptualize the active stance needed to dismantle racist power structures (Kendi, 2019). The adoption of anti-racism is closely tied to the development of *critical consciousness*, which applies more broadly to the development of awareness and active engagement in dismantling oppressive power structures (Freire, 1970).



Next, racialization, or “the extension of racial meaning to a previously racially unclassified relationship, social practice, or group” (Omi & Winant, 1986, p. 111), creates racial groups that are positioned to benefit white supremacy. In the early childhood context this impacts both teachers and children. The racialization of children by adults begins at a young age (Brown et al., 2010; Omi & Winant, 1986). Teachers may be racialized by parents or administrators and teachers may racialize young children and this impact the way that teachers interact with children, therefore may be an important component of classroom quality. For example, adultification, inaccurately seeing children as older or more mature than they are, of Black children can result in adults seeing Black boy’s developmentally appropriate behavior as violent and seeing Black girls as in need of less support and care (Epstein et al., 2017). Language is racialized as well. For example, being bilingual is viewed as a strength and sought-after skill by White parents but is seen as a deficit for Latine children (Shuck, 2006). For Latine children and DLLs, this racialization of language can lead to inequitable access to bilingual education and an outsized emphasis on English immersion (Chávez-Moreno, 2021). Asian and Multiracial children are almost entirely erased in this literature despite being the fastest growing demographic groups nationwide (Jones et al., 2021). White children are also racialized as innocent and too young for discussions of equity and race, as evidenced by the efforts of White adults to remove texts on race that may make White children uncomfortable (Feagin & Van Ausdale, 2001). These foundations of historical and present-day racism in ECE set the stage for the current studies. The following theoretical approaches incorporate these definitions for the dissertation studies.

### **Theoretical Perspectives**

## **Anti-Racist Early Childhood Education**

Though anti-racist theory has a long history in education, particularly in critical pedagogy (i.e., Paulo Freire, bell hooks), critical race (i.e., Gloria Ladson-Billings, David Gillborn, Daniel Solórzano), and postmodern (i.e., Judith Butler, Michel Foucault) theoretical approaches, application in the early childhood space is relatively recent (Escayg et al., 2020; Lynch et al., 2017). *Anti-racist early childhood education theory* is a theoretical perspective that highlights the centrality of racism, particularly in early childhood education, and proposes addressing racism throughout early childhood by focusing on dismantling systemic racism and white supremacy (Escayg, 2020). Anti-racist early childhood education centers on the significance of marginalized knowledge, experience, and resistance of white supremacy. It is rooted in a history of racism in the U.S., specifically as applied to the context of early childhood. Two key tenets of anti-racist early childhood education theory as proposed by Escayg (2020) are 1) interrogating early childhood education foundational knowledge through an anti-racist lens and 2) unmasking the white racial frame in pedagogical practices.

Specifically, Escayg (2020) highlights the ways that overwhelmingly White developmentalists have asserted authority in the content, empirical approaches, and practice in the ECE space. Through narrow definitions of developmentally appropriate practice and classroom quality based on ways of knowing rooted in whiteness, these White developmentalists perpetuate white power and privilege through the marginalization of racialized groups' ways of knowing. Additionally, quantitative approaches that are rooted in positivism and false notions of objectivity, which have historically been weaponized by white supremacy, are often privileged in education

policy spaces (Gillborn, 2018). In this study the CLASS may be considered a measure of classroom quality rooted in whiteness, therefore, it is important to critique the measure while simultaneously working to make sure it functions as equitably as possible, given its embeddedness in research and policy. Though the critique made by anti-racist ECE theory is useful and in itself resists white supremacy, anti-racism requires *praxis*, or action and reflection in equal measure (Casey, 2016). As such, anti-racist ECE theory operationalizes the reflection component through critique and unpacking of current operations of racism in early childhood, however, it also must be paired with an action-oriented perspective, such as that found within culturally relevant pedagogy.

### **Culturally Relevant Pedagogy**

*Culturally relevant pedagogy* (CRP) highlights the centrality of culture and racism in the classroom context, particularly in the relationships between teachers and students. CRP is an approach to teaching that centers students' cultural backgrounds in teacher-child interactions, and also actively challenges racism through fostering the development of critical consciousness in teachers and children. Underlying CRP are two essential themes: 1) care and compassion for students and 2) sociocultural or critical consciousness. Teacher's practice of CRP in the classroom emphasizes high academic expectations, the formation of positive cultural competence, and development of critical consciousness that prepares students to critique and challenge sociopolitical historical racism and inequity (Ladson-Billings, 1995; Morrison et al., 2008).

Most teachers choose the profession as a result of a desire to help students, however without the lens of critical consciousness, this help may be misguided or even

harmful in the case of White teachers who may unintentionally fall into the harmful trope of white saviorism (Casey, 2016). CRP as critically informed care and compassion for students, is an essential component of high-quality education that is not captured by measures of classroom quality such as the CLASS. The theoretical lens of CRP highlights the importance of understanding the impact of racism (an essential aspect of critical consciousness) on the interactions between teachers and students. In the absence of specific measures of critically conscious classroom quality, CRP suggests that the sociocultural context of the classroom, including the identities of teachers and children are important factors to be considered as important factors associated with classroom quality and school readiness. Further, it guides the research questions and the interpretation of findings centered in these themes. For the purposes of this dissertation, underpinnings articulated by CRP and specific principles proposed by anti-racist early childhood education serve to guide the research questions of the two following studies.

### **Present Studies Overview**

The aim of study 1 is to establish measurement invariance of the CLASS across Head Start teachers (Black, Latine, and White, due to sample size constraints). This is essential for understanding whether the measure is biased by teacher race/ethnicity. Given the high-stakes nature of the CLASS for Head Start teachers, it is essential that the measure function equivalently for all teachers, minimally at the level of intercepts, to compare means (Perlman et al., 2016; Thompson & Green, 2006). In line with anti-racist early childhood education, this study aims to critique and solution-build through the empirical investigation of the measurement functioning of the CLASS across teacher race/ethnicity. These findings will allow for the identification of potentially non-invariant

or biased parameters that may need to be accounted for by researchers using the measure, or specific items that may need to be interpreted with caution when they are used in policy and practice spaces.

For study 2, I proposed to examine profiles of classroom quality, dual language learner classroom composition as a predictor of quality profile membership, and profiles and demographic characteristics as predictors of child school readiness. Generally, the study investigated patterns of classroom quality, classroom sociocultural context (including children's racial and linguistic backgrounds and classroom language composition), and child school readiness. Specifically, the first aim was to establish profiles of classroom quality using the CLASS in a nationally representative Head Start sample. The second aim was to examine if the proportion of dual language learners (DLLs) in a classroom predicts the likelihood of a specific classroom profile (i.e., are classrooms with more DLLs more likely to be in a lower classroom quality profile?). Finally, the third aim was to examine the profiles of classroom quality, teacher-child racial ethnic match, language match, and child-level demographic characteristics in relation to children's school readiness, specifically social skills and approaches to learning.

These findings may have implications for a broader understanding of the heterogeneity in patterns of quality, in the context of racial and linguistic diversity. Further it may provide useful information for interventions related to teacher training and professional development on classroom quality and implications for how the heterogeneity in patterns of classroom quality might relate to school readiness over and above children's demographic characteristics and teacher-child racial/ethnic and

linguistic match. Finally, examining the relation of Head Start classroom dual language learner composition to quality profiles may shed light on DLL's differential access to certain types of quality.

There is an urgent need for data- and research-informed decision making in regard to the impact of racism in early childhood education, as there are clear disparities in the early childhood workforce and for children of color in ECE, created by white supremacist structures within education. The aim of these two studies was to explore these relations empirically and to provide actionable information regarding the nexus of racial/ethnic characteristics in educational settings, quality, and child school readiness outcomes. In the following sections, I detail the two proposed studies.

## CHAPTER 2

### STUDY 1 INTRODUCTION

#### **Classroom Assessment Scoring System Measurement Invariance Across Head Start Teacher Race/Ethnicity**

Quality is frequently focused on by researchers and policy makers as a point of intervention to improve the quality of early childhood education (ECE) and therefore maximize the positive benefits of ECE. There are two common definitions of classroom quality: process-oriented and structural quality (Pianta et al., 2005). The CLASS is an observational measure of process-oriented classroom quality that focuses on teacher-child interactions and is particularly important in ECE policy, research, and practice (Delaney & Krepps, 2021; La Paro & Pianta, 2003; Perlman et al., 2016). Specific to this study, classroom quality is largely defined and measured according to white middle class norms (e.g., behavioral manifestations of teacher sensitivity and regard for student perspectives; communication styles defined as high-quality in language modeling; behavior that would constitute encouragement and affirmation for quality feedback; Delaney & Krepps, 2021). The CLASS is used for accountability purposes in Head Start and many state quality rating and improvement systems (QRIS), often tied to ECE funding, directly impacting ECE teachers and the workforce more broadly (Office of Head Start, 2020; Quality Compendium, 2022). Additionally, the CLASS is used for individual evaluation of teachers, which may impact their job security. In this context, the CLASS has also become embedded in ECE technical assistance with the aim of improving teacher practice (Perlman et al., 2016; Pianta et al., 2008b). Therefore the measure plays an important role in the lives of the ECE workforce.

The ECE workforce is underpaid, with the workforce qualifying for public assistance in many states, and overworked (McLean et al, 2021). Further, research on ECE workforce inequities suggest pay disparities between White teachers and teachers of color and underrepresentation of teachers of color in lead teacher and director roles (McLean et al., 2021). These inequities across workforce race/ethnicity, the importance of classroom quality, and the widespread use of the CLASS measure highlight the importance of empirically investigating measurement invariance of the CLASS across teacher race/ethnicity. It is important that the measure is invariant across groups in order to draw valid and reliable conclusions (Knight et al., 2009). However, the CLASS does not address educational equity and only two studies of measurement invariance of the CLASS, neither focused on teachers, have been published to my knowledge (Downer et al., 2012; Thorpe et al., 2021). Despite the lack of studies on invariance the CLASS continues to be used to compare across and evaluate teachers. In this study, I aim to address this gap in the literature by investigating the measurement invariance of the CLASS across teacher race/ethnicity. In the following section, I first discuss the theories underlying this study, then discuss relevant context and research literature, and propose the current study.

### **Theories**

Two theories frame the present study. First, anti-racist ECE theory, proposed by Escayg (2020) highlights the importance of addressing racism in ECE. Inequities in the ECE space take many forms, but stem from the broader societal structure of the white supremacist capitalist patriarchy (hooks, 1994). Anti-racist ECE synthesizes the ideas of critical and poststructuralist theorists and highlights this structure and the way that it



manifests in ECE. An anti-racist ECE perspective allows for simultaneous critique of these embedded white middle class norms and solution building, as proposed in this study. It is important to take steps to make the CLASS more equitable due to its embeddedness in the early childhood space. In this case, testing for measurement invariance across teacher race/ethnicity is an initial step in this direction.

Next, culturally relevant pedagogy frames this study in relation to critiquing and understanding whether the conceptualization of classroom quality is culturally responsive. The CLASS tool is a measure that assumes that the classroom quality is global and that there are certain aspects of teacher-child interactional quality that are beneficial to all children. However, culturally relevant pedagogy assumes that cultural responsiveness is the baseline for quality care (Ladson-Billings, 1995). In this context, the tenets of care and compassion for students and sociocultural consciousness form the foundation of pedagogy and teacher-child interactions (Ladson-Billings, 1995). For example, in the case of DLLs, support for home language in the ECE classroom benefits not only children's expressive Spanish vocabulary, but also their quantitative reasoning skills (Downer et al., 2012; Partika et al., 2021). Teachers can also be agents of positive racial/ethnic socialization and examine the concepts of fairness and justice with young children (Farago et al., 2015; Farago et al., 2019; Feagin & Van Ausdale, 2001). This is particularly critical as children in pre-k begin to develop perspective-taking and empathy toward the ages of four and five (Eisenberg et al., 2014; Flavell et al., 1981; Hinnant & O'Brien, 2007; Hoffman, 2001). However, culturally responsive quality is not measured by the CLASS and therefore teachers are not getting credit for this high-quality care through the CLASS.

## **Policy Context**

The policy context of the CLASS is central to understanding its importance and widespread use. During the most recent reauthorization of Head Start in 2007, there were large changes made to the conceptions of quality, including increasing teacher education requirements and requiring that Head Start monitoring processes use a valid and reliable observational tool to assess classroom quality and teacher-child interactions, such as the Classroom Assessment Scoring System (CLASS). Additionally, it implemented a Designation Renewal System in which underperforming Head Start grantees, as identified by their performance on a valid and reliable quality measure, were required to re-compete for funding after their five-year designation (Improving Head Start for School Readiness Act of 2007). According to the Office of Head Start under the Administration for Children and Families, the CLASS was the only tool that met the requirement mentioned by law in the Head Start renewal in 2007 for a valid and reliable observational tool to assess quality and teacher-child interactions and has since been ingrained in federal regulations as of fall 2020 (Office of Head Start, 2020; Administration of Children and Families, 2020). The Act introduced the need for the CLASS measure and resulted in its codification into federal regulation, underscoring the centrality of this particular measure of classroom quality to Head Start policy.

## **Workforce Inequity**

The ECE workforce is a group of caring and dedicated professionals that provide essential care to young children. However, it is rife with disparities and providers are underpaid and overworked. The median pay for preschool teachers in 2019 was \$14.67/per hour, which in many states qualifies ECE providers for public assistance (U.S.

Department of Labor; McLean et al., 2021). There are also racial/ethnic disparities in pay and leadership representation within the ECE workforce. On average after accounting for education status, Black ECE providers make 78 cents less than White providers. This disparity is even more stark in the preschool context, where Black teachers make \$1.71 less than White teachers (Austin et al., 2019; McLean et al., 2019). For Head Start providers, pay ranges widely by location, but the national average pay is \$31,579 annually (ZipRecruiter, 2022). Additionally, teachers of color are underrepresented and White teachers are overrepresented in ECE leadership positions, including lead teacher and director roles (McLean et al., 2021). Contextual factors may also lead to additional job stress. For example, Barajas-Gonzalez (2021) found that immigration hostility between 2016-2020 exacerbated the stress of Latina ECE providers. Further, the COVID-19 pandemic has put additional weight on a child care system that was already under strain and disproportionately impacted (and continues to impact) Black, Latine, Asian, Pacific Islander, and Indigenous communities (Austin, 2021; Schilder & Sandstrom, 2021; Yip, 2021). These challenges can lead to high rates of workforce burnout and turnover, which is estimated to be anywhere from 13-40%, depending on a host of factors (Farewell et al., 2021; Totenahgen et al., 2016). High rates of turnover and job stress experienced by providers not only negatively impact providers themselves, but also negatively impact the quality of ECE (Cumming, 2017; Grant et al., 2018; Kwon et al., 2019).

Despite these challenges, the ECE workforce is still expected to continually improve the quality of ECE. Accountability is central to ensuring an equitable education system for children and has resulted in changes in Head Start teachers reported practices

(Walter & Lippard, 2017). However, implementation of accountability policies can also have unintended consequences. For example, a 2021 ethnographic study of seven Head Start teachers and two program leaders—the first of its kind—found that these providers viewed the CLASS as a useful tool for continuous quality improvement purposes during the 2010-2013 era when it was only used for quality improvement. However, since the tool has been transitioned to a high-stakes instrument under the designation renewal system, the providers now associated it with stress and found tension between what the teachers thought the observers wanted to see and what the children needed during observations (Delaney & Krepps, 2021). Additionally, two Black providers noted that the CLASS promoted an understanding of quality rooted in white middle-class approaches to teaching. They voiced a strong concern about the cultural mismatch between the tool and their cultural experiences in the community and observers' ability to adequately capture culturally relevant high-quality interactions. Moreover, the high-stakes importance of the CLASS did not allow space for culturally relevant practices (Delaney & Krepps, 2021).

Particularly for providers in Head Start settings or programs in QRIS, poor performance on quality measures—specifically on the CLASS in the case of Head Start—can jeopardize funding for their job or center. These high stakes highlight why it is essential that the CLASS is assessing the same construct equivalently and without systematic bias across teacher/race ethnicity. This is important so that the CLASS is not contributing to or creating additional disparities among the ECE workforce. The attempt to improve accountability is important for the equitable care of children, but should not come at the cost of inequitable and dehumanizing treatment of the ECE workforce.

### **Equity and Measurement Invariance**

Historically, science has been rife with bias and has weaponized assumed objectivity to maintain existing power structures. For example, at the turn of the 19th century, scientific racism about the superiority of White people and the inferiority of people of color, particularly Black people, was used to justify institutional racism, segregation, and Jim Crow laws under the guise of objective science (Jackson et al., 2005). More recently, scientific texts such as the Bell Curve have been used to argue the inferiority of people of color (Dennis, 1995; Winston, 2020). Notably, in the late 1980s and early 1990s, many of the measures developed in social science were created by White researchers and validated with White samples and then assumed to be universal and generalizable to Black, Indigenous, and other people of color. Underlying this use of measures with samples of various racial/ethnic groups is the assumption that these measures were capturing the same construct across groups without systematic bias (i.e., measurement invariance) (Han et al., 2019). However, more recent statistical and theoretical work, particularly by scholars of color, has emphasized the importance of testing—rather than assuming—measurement invariance as an underlying prerequisite for cultural validity (Han et al., 2019). Measurement invariance involves using statistical analyses to test the assumption that a measure is capturing the same phenomena across groups and is a best practice in valid and reliable measurement (Knight et al., 2009). The CLASS was validated in a large sample across multiple states and has been used with various racial/ethnic groups. However, measurement invariance has not been tested across teacher race/ethnicity and it has therefore been assumed rather than empirically assessed. As such, testing measurement invariance and examining potential bias by teacher race/ethnicity is an essential step toward anti-racist research in early childhood

education.

## **CLASS Literature**

As discussed previously, the term classroom quality may be defined by either structural or process characteristics (Bayly et al., 2021; Valentino, 2018). For the purposes of this study, classroom quality refers to the quality of processes and interactions between teachers and children as measured by the CLASS Pre-K (La Paro et al., 2004; Pianta et al., 2015). There are many iterations of the CLASS that emphasize different developmental periods, from infant and toddler classrooms up to high school (e.g., Allen et al., 2015; La Paro et al., 2014). It measures three domains of teacher-child interactions: 1) emotional support, 2) classroom organization, and 3) instructional support based on the teaching-as-interactions framework (Hamre et al., 2013; Pianta et al., 2015).

### **Three Domains of CLASSroom Quality**

First, emotional support is rooted in attachment theory, which suggests that positive and warm relationships between a secure adult and a child and safety and predictability allow the child to feel safe enough to explore, learn, and grow (Bowlby & Ainsworth, 2013; Hamre et al., 2013; Perlman et al., 2016). It includes four areas: 1) positive climate, 2) negative climate, 3) teacher sensitivity, and 4) regard for student perspectives (Pianta et al., 2015). The emotional support domain has been shown to be associated with children's social competence (Burchinal et al., 2010; Curby et al., 2009; Downer et al., 2012; Mashburn et al., 2008), executive functioning (Weiland et al., 2013), and reading and math skills (Pianta et al., 2008), and buffers the relation between problem behaviors and approaches to learning (Dominguez et al., 2011).

Second, classroom organization focuses on the physical set-up and components of the classroom and the ability of educators to structure the learning environment to minimize distraction and problem behaviors. This includes behavior management, productivity, and instructional learning formats. Classroom organization has been positively associated with language and literacy skills and social competence (Keys et al., 2013; Xu et al., 2014).

Finally, instructional support is rooted in a Vygotskian theoretical perspective, which proposes that children build knowledge through scaffolding and captures educators' ability to support this knowledge and growth through instructional practices (Fernyhough, 2008; Hamre et al., 2013; Perlman et al., 2016). This includes concept development, quality of feedback, and language modeling. Concept development has been positively associated with children's receptive vocabulary and applied problems abilities (Curby et al., 2009).

### **Classroom Quality and Child Outcomes**

The CLASS is widely used in early childhood research and the tool has been used with many different racial/ethnic groups, dual language learners, and children with disabilities in the United States, as well as internationally (Carr et al., 2019; Hamre et al., 2014; Hu et al., 2017; Leyva et al., 2015; Limlingan et al., 2019; Sabol et al., 2013). Despite its widespread use, findings about its relation to children's outcomes are mixed. According to Teachstone (2017), the parent company and publisher of the CLASS measure, more than 200 studies have investigated the association between the CLASS and children's outcomes across the toddler to high school educational contexts. Indeed, there is a large body of work demonstrating significant positive associations between

components of the CLASS and particular child outcomes as mentioned previously. On the other hand, there are many studies with small or null findings—raising concerns about the predictive validity of the CLASS (Burchinal et al, 2011; Guerrero-Rosada et al., 2021; Mashburn et al., 2008; Weiland et al., 2013). For instance, two meta-analyses, each including large national early childhood datasets, found no relation between the CLASS and child outcomes until classroom quality reached a high threshold (Burchinal et al., 2011a; Burchinal et al., 2016). For the measure to have strong predictive validity, it would be negatively associated with child outcomes at low levels of quality, but instead no association has been found.

Additionally, Perlman and colleagues (2016) conducted a systematic review (35 studies) and meta-analysis (14 of 35 studies) examining the relation between pre-k classroom quality as measured by the CLASS and children's outcomes. They found evidence of two small, but significant relations between the CLASS and children's outcomes. Specifically, the classroom organization domain was associated with pencil tapping and instructional support was associated with social skills (pooled correlations .06 and .09, respectively). Of the studies included in the review, 16 studies were from the National Center for Early Development & Learning, which was a multi-state longitudinal study focusing on pre-kindergarten and served as a validation study for the CLASS (Perlman et al., 2016; Teachstone, 2017). These findings suggest a need for caution in acting on any individual study finding and raise concerns about the practical significance of the 200 studies mentioned by Teachstone examining the CLASS and individual child outcomes.



There may be many issues at play in the mixed findings in the literature. For example, relationships between the CLASS and child outcomes may be nonlinear, or dosage of high-quality classroom experiences (Zaslow et al., 2010), the length of assessment (Teachstone, n.d.), or measurement issues may be at play interfering with the predictive validity of the measure (Li et al., 2019). To untangle these mixed findings, more recent studies have taken various innovative approaches beyond examining linear relationships between quality and child outcomes, including conducting threshold analyses (Burchinal et al., 2016; Hatfield et al., 2016; Weiland et al., 2013), pattern-centered analytic approaches (Iruka & Morgan, 2014; LoCasale-Crouch et al. 2007;), and additional measurement investigations (Li et al., 2020; Styk et al., 2021). Some threshold analyses suggest that there is no relation between quality and child outcomes until quality reaches a high level (Burchinal et al., 2016; Hatfield et al., 2016), while Guerrero-Rosada and colleagues (2021) continue to find nonsignificant relations between the CLASS and child outcomes even at high levels of quality.

### **CLASS Measurement**

A precursor to assessing measurement invariance is to establish a well-fitting factor structure. Various studies have examined the measurement properties of the CLASS. According to the developers of the CLASS, the proposed measurement model includes three factors indicated by 10 items (Pianta et al., 2015). A recent meta-analysis of the CLASS factor structure, which included data from 26 studies using the CLASS, provided support for the three-factor model originally proposed by the developers (Li et al., 2019). On the other hand, eight separate studies have been conducted examining the measurement model of the CLASS, and these individual studies have found poor to

moderate fit of the three-factor model. In addition to the proposed three factor model, other studies have examined one-, two-, and bifactor models. For example, Sandilos and colleagues (2014) found that a bifactor model fit the data best, but still did not fit the data well in a sample of North Carolina and Pennsylvania preschoolers. Further, Towson and colleagues (2016) found that the three-factor model fit their data better than one-, two-, or bi-factor models, but still did not fit the data well. The CLASS has also been utilized internationally and confirmatory factor analyses have been conducted to examine the factor structure of the measure in various contexts, including Finland, Germany, Chile, and China. Across these studies, the three-factor model did not fit the data well but fit better than the alternate models (Hu et al., 2016; Leyva et al., 2015; Pakarinen et al., 2010; Von Suchodoletz et al., 2014).

Across the eight studies examining the factor structure of the CLASS in preschool, pre-k, or kindergarten, generally the three-factor structure produced mixed global fit statistics (Li et al., 2019). The CFIs ranged from .84-.93, which is lower than recommended  $>.95$  recommended for good fit; however, four studies reported CFIs above .90, meeting the criteria for adequate fit (Hamre et al., 2014; Hu & Bentler, 1999; Hu et al., 2016; Towson et al., 2016; Von Suchodoletz et al., 2014). The RMSEA was higher than the recommended  $<.08$  for adequate fit of the 3-factor model across all studies in the meta-analysis (RMSEA = .11-.23; Browne & Cudeck, 1993; Hamre et al., 2014; Hu et al., 2016; Leyva et al., 2015; Pakarinen et al., 2010; Sandilos et al., 2014; Towson et al., 2016; Von Suchodoletz et al., 2014), with the exception of Hamre et al. (2013; RMSEA = .05). Finally, the SRMR ranged from .05-.09, which met the recommended  $<.10$  criteria for adequate fit (Hamre et al., 2013; Hamre et al., 2014; Hu

& Bentler, 1999; Hu et al., 2016; Leyva et al., 2015; Pakarinen et al., 2010; Sandilos et al., 2014; Towson et al., 2016; Von Suchodoletz et al., 2014; ). These studies demonstrate the mixed findings in relation to the CLASS factor structure and the recommended three-factor model.

Two studies have moved beyond establishing a factor structure to examine measurement invariance of the CLASS. First, Downer and colleagues (2012) examined measurement invariance of the CLASS in pre-k classrooms with varying composition of dual language learners and Latine children. They found evidence of strong measurement invariance of the CLASS across the dual language learner and Latine children composition using the National Center for Early Development and Learning's (NCELD) Multi-State Study of Pre-Kindergarten (Multi-State Study) and State-Wide Early Education Programs Study (SWEEP Study). Second, Thorpe and colleagues (2020) examined longitudinal invariance of the CLASS in a sample of Australian pre-k (3- and 4-year-old) and year 2 (age 7-8-year-old) children and found that the CLASS was longitudinally invariant across time, with time defined as varying numbers of observation cycles. Downer and colleagues (2012) conducted measurement invariance using a three factor model, whereas Thorpe and colleagues (2020) conducted measurement invariance across time for each of the three factors separately (i.e., assessed longitudinal invariance of emotional support across time, classroom organization across time, and instructional support across time). The measurement invariance of the CLASS across teacher race/ethnicity remains an open and essential question.

## **Present Study**

Classroom quality as measured by the CLASS tool plays an important role in early childhood policy, research, and practice. Racism plays a central role in U.S. society, including in research and education. As such, empirically investigating invariance across teacher race/ethnicity of the measure is essential to achieving equitable policy and practice in ECE that supports fair treatment of ECE teachers. This study addressed the following research question:

**Research Question 1:** Does measurement invariance for the CLASS hold across teacher race/ethnicity (Black, Latine, White)?

These racial/ethnic groups were selected due to sample size constraints. It is important to examine invariance across American Indian/Alaska Native teachers, Asian and Pacific Islander teachers, and teachers of other races in the future. Further, it is important to note that analyzing invariance across these groups, analysis inherently essentializes them as monolithic, although there is heterogeneity within, as well as between, these groups.

These are two noted limitations of the current study.

## CHAPTER 3

### STUDY 1 METHOD

#### **Participants**

Participants were drawn from the Head Start Family and Child Experiences Survey (FACES) 2009 cohort. The study used a probability sampling method to select programs, centers, and classrooms and used equal-probability sampling at the child level. Data for the present study was drawn from the second wave of data collection, which was collected during the spring of 2010 ( $n = 486$  classrooms). Observational data was only collected at T2 and was organized by the classroom lead teacher. Teachers in the study were almost exclusively female and were 35.7% White teachers, 28.9% Black teachers, 19.5% Hispanic/Latine teachers, 12.1% teachers who identified as “Another Race,” 2.0% Asian or Pacific Islander teachers, and 1.8% American Indian/Alaska Native teachers.

#### **Procedures**

Observations were conducted during the spring of 2010 (T2) by trained observers and were designed to be minimally intrusive to the classroom environment. Observers spent at least 4 hours in classrooms and conducted observations in the morning using paper documents to record observations that were later transferred to the computer (Malone et al., 2013).

#### **Measures**

##### *Classroom Assessment Scoring System*

The classroom assessment scoring system (CLASS; Pianta et al., 2008) was used to assess the interactional aspects of classroom quality in Head Start classrooms. This

observational measure comprises 11 areas, grouped into three domains: emotional support (e.g., teacher sensitivity), classroom organization (e.g., learning formats), and instructional support (e.g., language modeling). Observations occurred over four cycles and ranged from 1 “minimally characteristic” to 7 “highly characteristic.” The measure demonstrated high internal consistency, with Cronbach’s alphas ranging from .79-.81, and demonstrated evidence of good inter-rater reliability, with an average of 87% agreement (Malone et al., 2013).

### ***Demographics***

Demographic data for the teachers were gathered through the teacher survey. For this study, teachers provided their race/ethnicity. The response options included the following racial/ethnic groups: African American, White, Multiracial/Biracial, American Indian or Alaska Native, Asian or Pacific Islander, and other race and the option to identify as Hispanic or Non-Hispanic.

### **Analytic Plan**

#### ***Preliminary Analyses***

Descriptive statistics and correlations among study variables were conducted prior to the primary analyses. Specifically, means, standard deviations, skewness, kurtosis, and bivariate correlations were examined using SPSS.

#### ***Primary Analyses***

Measurement invariance analyses were conducted using Mplus version 8. To conduct the invariance analyses, I first conducted confirmatory factor analysis (CFA) to establish the measurement model of the CLASS in the current sample. The model as proposed by the publishers is a three-factor model, including emotional support indicated

by 4 variables, instructional support indicated by 3 variables, and classroom organization indicated by 3 variables. To scale the factor I used the Mplus default, which sets the first item of each factor loading to 1. Missing data was handled using full information maximum likelihood. Model fit was examined using local fit (factor loadings) and global model fit indices, including the chi-square, root mean square error of approximation (RMSEA), comparative fit index (CFI), and standardized root mean square residual (SRMR). Close fit of the model was determined if two of the following three criteria are met: non-significant chi-square,  $RMSEA \leq .08$  (Browne & Cudeck, 1993),  $CFI \geq .95$ , and  $SRMR \leq .08$  (Hu & Bentler, 1999). I also examined and tested modification indices that might improve the fit of the chi-square statistic.

Next, I conducted measurement invariance analyses across teacher race/ethnicity (Black, Latine, and White). Following Byrne's (2012) method of measurement invariance testing, I increasingly constrained parameters, first testing for configural invariance (constraining the structure of the model to be equal across groups), weak invariance (constraining the factor loadings to be equal between groups), and strong invariance (constraining the intercepts to be equal between groups). I identified whether the following changes between models were observed:  $CFI < .01$ ,  $RMSEA < .015$ ,  $SRMR < .03$  (Chen, 2007; Keum et al., 2018).

Partial strong invariance was tested in the case that the strong invariance model fits significantly poorer than the weak invariance model. Partial invariance uses modification indices to identify individual non-invariant parameters and free these parameters in a strong invariance model. Between-group inferences can be drawn from partial strong invariance models, as long as a minimum of one factor loading and one

intercept for each factor is equivalent between groups (Byrne et al., 1989; Thompson and Green, 2006).



## CHAPTER 4

### STUDY 1 RESULTS

#### **Preliminary Analyses**

First, descriptive statistics were analyzed. These included zero-order correlations, means, standard deviations, skewness, and kurtosis for all study variables. Correlations were in the expected magnitude and directions. Items within subscales (i.e., concept development and quality feedback) tended to be significantly positively correlated, with the exception of the negative climate item, which was negatively correlated with all items as expected (see Table 1). Kurtosis was high (i.e., outside of the recommended -2 and +2 range) for the emotional support composite and quality feedback item (see Table 2), suggesting nonnormality (West et al., 1995). To account for this nonnormality and missing data, a robust maximum likelihood estimator (MLR) was used in subsequent analyses.

#### **Primary Analyses**

##### *Confirmatory Factor Analysis*

The teaching-as-interactions model suggests that the CLASS is composed of three factors underlying 10 items, where all three factors are allowed to covary. First, I conducted confirmatory factor analysis to establish this three-factor model in the full teacher sample and to establish a well-fitting baseline model in the full sample, before establishing a well-fitting baseline model for each teacher racial/ethnic group. The global fit indices indicated that the model did not fit the data well (see Table 3). However, local fit indices demonstrated that the factor loadings were in the expected direction and the standardized factor loadings were medium large (.556-.940), indicating good

measurement quality (see Table 3; McNeish & Hancock, 2016). This pattern of findings, poor global fit indices with good local fit indices, suggested evidence of the reliability paradox (Hancock & Mueller, 2011; Kang et al., 2016; McNeish & Hancock, 2016; Meade et al., 2008). The reliability paradox refers to the phenomena that good measurement quality, as indicated by local fit (specifically factor loadings), is inversely related to good global fit, particularly with approximate fit indices such as the chi-square, RMSEA, and SRMR (Hancock & Mueller, 2011; McNeish & Hancock, 2016). In a simulation study examining the reliability paradox in the context of approximate fit indices in single-group CFA models, McNeish and Hancock (2016), found that under conditions of excellent measurement quality (i.e., factor loadings of 0.9), approximate fit indices produced what are typically considered “poor fit” indices (RMSEA  $\geq$  .20, SRMR  $\geq$  .14, CFI  $\leq$  .775). The opposite was also true, models with poorer measurement quality (i.e., factor loadings of 0.4) produced commonly supported “good fit” indices (RMSEA  $\geq$  .20, SRMR  $\geq$  .14, CFI  $\leq$  .775). Therefore McNeish and Hancock (2016) recommend interpreting model fit based on both measurement quality and a cautious interpretation of approximate fit indices. They recommend interpreting approximate fit indices in the context of measurement quality.

Kang and colleagues (2016) expanded this work to the multigroup CFA measurement invariance context. In a simulation study examining the reliability paradox and global fit indices in the context of multigroup CFA and measurement invariance testing, Kang and colleagues (2016) found that even the CFI, recommended by McNeish and Hancock (2016), was associated with the reliability paradox in the context of multiple group tests of invariance and therefore recommended interpreting this fit index

with caution. They found that the change in McDonald's Noncentrality Index ( $\Delta$  MNCI) was more robust to the issues associated with the reliability paradox when assessing measurement invariance. They found that the CFI varied significantly across quality of measurement, number of indicators, and sample size, but MNCI was resilient to these factors. MNCI is not offered in Mplus or other software packages and therefore was calculated by hand using the equation provided in the summary of fit indices equation by Hu and Bentler (1999). Moving forward, the change in MNCI ( $< .02$ ) and local fit were used as the primary fit indicators, and change in CFI ( $< .01$ ), RMSEA, and SRMR were used as supplementary indicators (Cheung & Rensvold, 2002).

### ***Measurement Invariance***

To conduct measurement invariance analyses, I first established a well-fitting baseline model for Black, Latine, and White teachers separately (see Table 3 for all global model fit indices). Next, I tested for configural invariance where the structure of the models is constrained to be equal across groups. These global fit indices indicated generally poor fit; however, in the context of the reliability paradox, I also examined the measurement quality (see Table 4). The factor loadings ranged from .556-.940, indicating medium-high measurement quality (except for negative climate), suggesting good alignment between the data and model (McNeish & Hancock, 2016). Therefore, I moved on to test for metric/weak measurement invariance, where the factor loadings are constrained to be equal across groups. The  $\Delta$  MNCI was -.004, which was below the recommended threshold of -.02; therefore, I moved on to test for scalar/strong invariance. Scalar/strong invariance involves constraining the intercepts across groups. The  $\Delta$  MNCI was -.025, which is slightly higher than the Cheung and Rensvold (2002) recommended

cutoff. Therefore, I examined modification indices to determine which parameters might be freed to improve the chi-square statistic and test for partial scalar invariance. The largest modification indices were two covariances in the Black teacher group. The modification indices indicated that freeing the covariance between the positive climate and concept development residuals would improve the chi-square statistic by 19.294 and freeing the covariance between the positive climate and language modeling residuals would improve the chi-square by 19.658. Of note, positive climate and concept development have been shown to be related in previous research (Howes et al., 2013). The fit of the model with this parameter freed and resulted in model fit with improvements in the CFI, RMSEA, and the same SRMR, indicating partial scalar invariance (see Table 3). Additionally, Latine teachers served as the reference group and White teacher's standardized factor means were Emotional Support = .203, Classroom Organization = .196, Instructional Support = .437. Relative to Latine teachers, Black teacher's standardized factor means were Emotional Support = -.526, Classroom Organization = -.460, Instructional Support = -.131.

## CHAPTER 5

### STUDY 1 DISCUSSION

#### **Summary of Findings**

In this study, I examined measurement invariance of the CLASS across teacher race/ethnicity (specifically Latine, Black, White). First, I tested the publisher recommended three-factor measurement model in the current sample. I found that the three-factor CLASS CFA model fit poorly according to approximate and global fit indices, but there was evidence of medium to high measurement quality, suggesting evidence of the reliability paradox (McNeish et al., 2016). Therefore, I added McDonald's NCI as an indicator of global fit and evaluated the global model fit indices in the context of local fit and moved on to test measurement invariance (McNeish et al., 2016; Kang et al., 2016). Then I tested the measurement invariance across teacher race/ethnicity which entailed establishing a baseline model for each racial/ethnic group (Latine, White, and Black due to sample size constraints), testing for configural, metric, scalar, and partial scalar invariance (Byrne, 2012). I found evidence of partial scalar invariance, indicating that there was evidence of invariance across teacher race/ethnicity at the level of the intercepts and that means may be compared across these groups (Byrne, 1989; Thompson & Green, 2006). However, the parameter that was freed to establish partial scalar invariance, the residual covariance between positive climate and concept development, should be interpreted with caution for Black teachers. In policy and practice this indicates that scores on these parameters should be interpreted with caution.

## **Explanation of Findings**

### **CFA and the Reliability Paradox**

In this study I first conducted a CFA to establish a measurement model in the current sample and found evidence of poor global fit across CFI, RMSEA, and SRMR. This finding is similar to the majority of research on the CLASS using CFA and SEM. With a few exceptions (i.e., Li et al., 2019), researchers have found evidence of poor global fit and this may provide a potential explanation of the finding in this study, at least in the case of the CLASS in Head Start (Hu et al., 2016; Leyva et al., 2015; Pakarinen et al., 2010; Sandilos et al., 2014; Towson et al., 2016; von Suchodoletz et al., 2014). However, the reliability paradox highlights the potential inverse relation between global and local fit, which is particularly problematic for global fit indices such as the chi-square, RMSEA, and SRMR, whereas the CFI has been found to be more robust to this issue (McNeish et al., 2016). This finding is born out in the CLASS measurement literature where the Li and colleagues (2019) meta-analysis highlighted that four CFA studies had adequate CFI fit (Hamre et al., 2014; Hu et al., 2016; Towson et al., 2016; Von Suchodoletz et al., 2014), but none of the included studies met the adequate global fit criteria for RMSEA and SRMR. A contribution of this study is applying an understanding of the reliability paradox to the measurement model of the CLASS. Future studies conducting CFAs should consider the global fit indices in the broader context of local fit and utilize fit indices that are less prone to this issue such as the CFI and for measurement invariance change in McDonald's NCI (McNeish et al., 2016; Kang et al., 2016).

### **Partial Scalar Measurement Invariance**

Measurement invariance across intercepts (scalar invariance) is a baseline step for mean comparison across groups. The finding of partial scalar invariance aligns with previous literature on CLASS measurement invariance. Previously, Downer and colleagues (2012) and Thorpe and colleagues (2020) found evidence of scalar invariance across classrooms with varying proportions of DLL composition and various proportions of Latine classroom composition as well as across time. Though in the current study I found evidence of partial scalar invariance, only one parameter was freed to achieve this level of invariance and the majority of parameters were invariant across groups. This is important to note because within partial invariance as the proportion of non-invariant parameters increases as does the potential bias (Chen, 2008). To achieve partial invariance, I freed the residual covariance between the concept development and positive climate for Black teachers. Empirically, freeing this parameter would result in the second-best improvement of fit according to modification indices. Additionally, it did not require any cross loading of items or freeing a parameter across groups, both of which would posit larger changes to the measurement model. Theoretically, it is reasonable to assume that the errors of concept development and positive climate may be related, as these variables have been found to be associated in the past and therefore may have associated errors (Howes et al., 2013).

**Research and Practical Considerations.** Findings of partial scalar invariance are very common, occurring in about one-third of measurement invariance testing publications in the psychological sciences (Putnik & Bornstein, 2016). However, researchers differ on how partial invariance should be handled in research as well as the clinical significance of partial invariance and its impact on the use of measures in

practice. Current recommendations for managing partial invariance include applying some form of sensitivity testing in which the statistical test of interest (e.g., mean comparison, regression, etc.) is conducted using a fully invariant (with invariant and non-invariant parameters constrained) and partially invariant model. These are then compared substantively to determine if there was a practice difference in the implications of the non-invariance (Chen, 2008). As applied to the current study, these recommendations from the measurement invariance literature suggest that researchers test the sensitivity of partial invariance in their analysis using the CLASS, paying particular attention to this sensitivity for Black teachers on concept development and positive climate. For Head Start center directors, teachers, and policy makers this is an important consideration for the use of the CLASS. This will be discussed further in the implications section.

### **Theoretical Interpretation**

The finding of partial scalar invariance is relatively positive, because it indicates that measurement bias on teacher race/ethnicity of the CLASS itself is likely not a large factor exacerbating workforce inequity. However, this is not to say that there may not be other sources of measurement bias or that bias on race/ethnicity does not exist for teacher of other racial/ethnic groups. This finding is not what was expected based on anti-racist ECE theory (Escayg, 2020). Anti-racist ECE theory would suggest that measures, such as the CLASS, that are widely used in early childhood spaces often prioritize the culture of power and the interaction styles embodied by white middle class values and therefore we would have expected a greater degree of non-invariance (Escayg, 2020). This finding is positive, particularly given the measure's importance in policy and practice contexts. However, there is still potential bias in *what* is being prioritized by the CLASS. The



definitions and operationalizations of the CLASS itself remain colorblind. Expanding the definition of quality to center equity remains central to social justice efforts in ECE. This could include incorporating a fourth factor to the CLASS model that operationalizes culturally responsive practice as a core tenet of quality. This way educators of color are getting credit for the high-quality instruction they are providing that is not currently captured by the CLASS (Delaney & Krepps, 2021).

### **Implications for Policy and Practice**

This study addressed a measurement focused research question. Though this may be thought of in terms of basic research, its findings have implications for policy and practice. The findings provides evidence of partial scalar measurement invariance of the CLASS across teacher race/ethnicity, which allows for the means to be compared across these groups. This finding is a baseline best practice for the use of measures with multiple racial/ethnic groups and groups with other marginalized identities (Knight et al., 2009). In policy contexts, this means that Head Start (ACF and HHS) and states that use the CLASS as part of QRIS can use the CLASS as a measure of classroom quality across White, Black, and Latine teachers and have reasonable confidence that the means of the CLASS are capturing the same construct across teacher race/ethnicity.

However, it is important to be clear that the findings of this study do not suggest that the CLASS is an anti-bias or equity focused measure. This means that although there was evidence of partial scalar measurement invariance on teacher race/ethnicity (for Black, Latine, and White teachers), it does not explicitly or intentionally introduce equity content. In practice this means that directors, coaches, and teachers may use CLASS scores to help inform areas of teacher strengths to build on and areas where there is room

for growth that may be addressed by tailored professional development. In policy making decisions regarding funding, for example, the designation renewal system, this study shows that the CLASS can be used as a guide or indicator of interactional quality but utilizing CLASS scores alone remains overly stringent and may have inequitable impact.

Another motivator of the study was the existing racial inequities that early childhood educators face. Across the ECE workforce, low wages, demanding work, and minimal benefits threaten the sustainability of the industry, yet the needs for high quality care remain (McLean et al., 2019). Within these difficult circumstances there are racial/ethnic inequities that leave women of color underrepresented in leadership positions (which come with better pay and benefits) and overrepresented in teacher aide roles (with the least compensation; McLean et al., 2019). The high stakes nature of the use of the CLASS tool is one additional stressor that can take an informative tool on teacher practice and turn it into a significant stressor that may feel as though it does not capture all of the aspects of quality that many teachers of color provide (Delaney & Krepps, 2021). Findings of partial scalar measurement invariance provide some comfort in that the measure is capturing the same construct across groups. However, this does not address the way the tool is being used and the resulting implications for equity in the ECE workforce.

Additionally, the parent company of the CLASS, Teachstone is making efforts to both change its existing measure and modify observation training for raters to better address equity concerns. They released CLASS PreK-3 version 2 which is designed to be more inclusive in terms of the age/grade composition of the measure which has important implications for assessors and credentials. Additionally, they have updated examples and

definitions within the existing framework to be more inclusive and have expanded definitions or operationalization of certain rated behaviors. Further, the second version of the CLASS appears to incorporate ideas of equity into the existing framework but does not appear to add equity as an explicit and unique component of quality. At this point in time (spring 2022) examples of changes have been provided, but the specific details have not been released. In the context of this study, this change means that findings of this study should be tested with the new measure once training is implemented, changes are rolled out, and data is collected.

### **Future Directions, Limitations, and Conclusions**

This study provides evidence of partial scalar invariance on the CLASS across teacher race/ethnicity, which is an important baseline step toward more equitable measurement, as it provides empirical support for the use of the tool to compare means across groups (Thompson & Green, 2006). Though this is meaningful, this study was not without limitations. First, this study only tested for invariance across Black, Latine, and White teachers. The measurement invariance of the CLASS should also be tested across Asian, Pacific Islander, Indigenous, Multiracial, and other racial/ethnic groups to examine measurement invariance across these racial/ethnic groups. Additionally, this should be tested with the CLASS version 2. Further, assessment of measurement invariance across observer characteristics is needed. Though observational tools may be helpful in preventing a precarious situation in which teachers are asked to self-report classroom quality, it introduces an additional human layer in between the measure and the data. The observer is subject to their own biases and perceptions and without knowing who the observer is or what social location they occupy, it is not possible to understand

the implications of this bias for measurement. For example, an observer's implicit biases may unintentionally affect their ratings. Therefore assessment of measurement invariance across observer race/ethnicity is needed to ensure that observers are capturing the same construct across observer race/ethnicity.

Further, statistical innovation in testing for invariance across intersectional identities such as race/ethnicity and language using approaches such as moderated nonlinear factor analysis is another important future direction. Another possibility could be to examine invariance across the interaction of teacher and observer race/ethnicity. Next, additional research should be done to look at the reliability paradox and the CLASS. Though the findings of this study are promising in regard to the ability of the CLASS to capture a similar construct across these groups, **this study does not address the fundamental lack of equity, anti-bias, or culturally responsive content** captured by the CLASS measure. Considering equity as a fundamental component of quality is essential and should be captured in widely used and influential measures such as the CLASS. The findings of this study provide promising evidence that in the existing framework posited by the class for interactional quality between teachers are functioning similarly at the level of intercepts across teacher race/ethnicity. The CLASS tool should be built on to expand equity.

## CHAPTER 6

### STUDY 2 INTRODUCTION

#### **Profiles of Classroom Quality, Equity Predictors, and School Readiness Outcomes**

Access to quality ECE is essential for educational equity in early childhood, but is not always equitable (Hardy et al., 2020). The patterns of quality that shape children's classroom experiences and school readiness outcomes are key variables for the effective investment of public resources and for determining paths forward in research. Unfortunately, historic and present-day racism permeates all aspects of life, including children's experiences of quality education (Hardy et al., 2020; Ladson-Billings, 2013). Children of color are more likely to attend under-resourced schools and have less access to highly trained and well compensated teachers (Bassok & Galdo, 2016; Morgan & Amerikaner, 2018). Head Start aims to alleviate these disparities through a variety of approaches, including high-quality federally funded preschool and wrap-around services for families. However, there is still variability in quality within Head Start (Morris et al., 2018). One approach to understand this variability in patterns of classroom quality is through the examination of latent profiles of quality in Head Start classrooms (Curby et al., 2009; LoCasale-Crouch et al., 2007). This approach goes beyond a variable-centered approach to center the way quality indicators pattern together in the classroom rather than averaging across indicators. However, quality profiles have yet to be explored from the lens of the sociocultural context of the classroom (i.e., teacher, child, and classroom composition racial/ethnic and language). Access to high-quality ECE, specifically

through the lens of various patterns of classroom quality, for children from marginalized backgrounds within Head Start is a central equity issue, particularly in relation to school readiness. In this study I aim to examine 1) profiles of classroom quality in a nationally representative Head Start sample, 2) the relation between quality profiles, sociocultural context, and school readiness, and 3) dual language learner (DLL) classroom composition as a predictor of these profiles. To situate the present study, I introduce in the following sections the theoretical perspectives guiding this study, relevant research literature on classroom quality profiles, school readiness, and demographic characteristics in ECE.

### **Theoretical Perspectives**

Developmentalists dominate the field of early childhood, which can be problematic for racial/ethnic equity. They tend to overemphasize commonalities and universalities among young children, which is a perspective that aligns with colorblind ideologies, and many are often reticent to consider potential implications of racism in the lives of young children (Escayg, 2020). From these perspectives, young children are often viewed as innocent and outside the operation of racism and other serious issues plaguing the adult world, precluding researchers from examining racism in early childhood (Feagin & Van Ausdale, 2001). From the perspective of *anti-racist ECE theory*, understanding the operation of racialized qualities—such as dual language learner classroom composition and teacher-child racial/ethnic and language match—is essential to understanding the patterns of quality for all children. Pairing this anti-racist perspective with an action-oriented lens, such as *culturally relevant pedagogy*, is

essential. These approaches allow for critique of dominant paradigms (like those demonstrated in the CLASS conceptualization of quality) and advocacy for the creation of improved measures, while simultaneously acknowledging the political and research centrality of the CLASS measure and using it in ways that consider the relevance of children's racial/ethnic and language context.

### **Classroom Quality Profiles Literature**

Classroom quality is central to researchers' understanding the impact of early childhood education and is a malleable target for policy makers to help to improve ECE and child outcomes (Perlman et al., 2016). Quality has been defined in a variety of ways across the research literature, but in this study, it is defined as teacher-child interactional quality as assessed by the CLASS, due to its importance to policy and practice, particularly in Head Start (Perlman et al., 2016). Previously, quality has primarily been studied from a variable-centered approach, examining predictors and outcomes of quality, emphasizing the variable itself, rather than the nuances and heterogeneity within classroom quality (Perlman et al., 2016). However, more recently, researchers have begun to examine patterns of classroom quality using a pattern-centered analysis approach, which allows for the empirical investigation of heterogeneous patterns of quality, and a closer empirical representation of quality as experienced by teachers and children in the classroom. This is a relatively new approach with only ten published early childhood studies using it to my knowledge, and only two focusing specifically on Head Start samples (Bulotsky-Shearer et al., 2012; Lippard et al., 2019). Next, I review the

content of these studies to communicate the state of the current research in this area, and to highlight the gaps in this literature that will be addressed by the current study.

Researchers have taken different approaches to examining the heterogeneity of classroom quality (i.e., patterns or profiles of classroom quality), the early childhood populations studied, the sociocultural context of the classrooms, and the association of quality profiles with school readiness. First, in their approaches to examining heterogeneity in patterns of quality some have used indicators from a single measure of classroom quality (e.g., Curby et al., 2009; LoCasale-Crouch et al., 2007) and others have used a host of different quality measures to construct the profiles (e.g., Iruka & Morgan, 2014; Lippard et al., 2019). These approaches serve different purposes. Using indicators from a single measure of classroom quality to create profiles allows for a more focused examination of indicators within a particular definition of quality. On the other hand, including multiple measures of quality emphasizes breadth and may provide insight into how differing definitions or operationalizations of quality pattern together. In this study I use a single measure indicator of quality profiles, the CLASS, to prioritize more direct implications for policy and practice.

Researchers have also focused on different early childhood populations when examining these profiles with some studying state pre-k (Curby et al., 2009; LoCasale-Crouch et al., 2007), utilizing large national datasets such as the Early Childhood Longitudinal Study- Kindergarten cohort (Iruka & Morgan, 2014), international samples (Hoang et al., 2019 [Vietnam]; Salminen et al., 2012 [Finland]; Salminen et al., 2018



[Finland]), and others examining large nationally representative (Bulotsky-Shearer et al., 2012; Lippard et al., 2019) or local Head Start samples (Lee & Bierman, 2016). However, the use of a large nationally representative Head Start sample using the CLASS indicators to create profiles has yet to be studied. Additionally, the relation between quality profiles and child outcomes has been examined in a variety of ways. Most studies examined both socioemotional and academic child outcomes (Bulotsky-Shearer et al., 2012; Curby et al., 2009; Hoang et al., 2019; Lee & Bierman, 2016; Salminen et al., 2018), while others focused only on academic outcomes (Iruka & Morgan, 2014; Lippard et al., 2019). These studies produced mixed findings, however, Lee & Bierman (2016) and Curby and colleagues (2009) found significant associations between profiles of quality and children's social skills. Next, most studies of quality profiles did not address sociocultural context (Hoang et al., 2019; Lippard et al., 2019; LoCasale-Crouch et al., 2007; Salminen et al., 2012; Salminen et al., 2018). However, Curby and colleagues (2009) included child demographic characteristics and Iruka and Morgan (2014) limited their study to Black children. Finally, others took a different approach by incorporating contextual factors as indicators of the profiles themselves (Bulotsky-Shearer et al., 2012; Lee & Bierman, 2016). A more comprehensive approach to sociocultural classroom quality that specifically examines teacher-child match, classroom composition, and child demographics as predictors of quality in addition to profiles is currently missing from the literature.

Studies also differed in the number of quality profiles identified. Four studies found 3 profiles (Hoang et al., 2019; Iruka & Morgan, 2014; Lippard et al., 2019;

Salminen et al., 2018), two found 4 profiles (Lee & Bierman, 2016; Salminen et al., 2012), two found 5 profiles (Curby et al., 2009; LoCasale-Crouch et al., 2007), and one found 6 profiles (Bulotsky-Shearer et al., 2012). However, it is important to note that nearly all of these studies used different indicators for the profiles, making it difficult to compare the number of profiles found across studies. For example, LoCasale-Crouch and colleagues (2007) examined heterogeneity in quality in state pre-kindergarten classrooms and used the CLASS domains to estimate quality profiles and found evidence of five profiles of quality using a cluster analysis, one of which was generally high quality, and one was generally low, with the remainder falling in between. Curby and colleagues (2009) extended this work to predict children's school readiness outcomes from these profiles in state pre-kindergarten. They found that children in the profile with high concept development, the classrooms in which teachers highly prioritized problem solving and application, showed the greatest gains in vocabulary and problem solving, relative to those in the profile with the highest overall quality (Curby et al., 2009; La Paro et al., 2004).

These findings highlight the need to understand the arrangement of components of quality that are essential to improving child outcomes, particularly school readiness. To summarize a study using a nationally representative Head Start sample, examining classroom quality profiles as indicated by the CLASS, examining positive social skills (i.e., cooperation) and approaches to learning, and more holistically conceptualizing classroom sociocultural context is needed. For example, a classroom quality profile in which a teacher is highly sensitive to children's needs and skilled in behavior

management, but low in concept development may positively predict children's social skills. Or on the other hand, a classroom quality profile in which a teacher is high in concept development and in behavior management, but low in sensitivity may positively predict children's approaches to learning. Additionally, the sociocultural context of the classroom is another theoretically important avenue of research.

### **Sociocultural Context of Quality**

The sociocultural context of Head Start, teachers, classrooms, and children influence the experiences that children have in early childhood. Within the context of the low-income and working-class children and families that Head Start serves, there are additional racialized contexts, including teacher and child race/ethnicity, language, the match between teachers and their students on these characteristics, and the composition of classrooms. Though the study of classroom quality profiles is relatively new, and therefore research regarding quality profiles and sociocultural context is limited, there is a wealth of literature examining classroom quality and sociocultural context from a variable-centered approach. Below I provide an overview of the literature examining aspects of the sociocultural context of classroom quality as captured by the CLASS to frame this study.

#### **Child Demographics and Classroom Quality**

##### ***Child Demographics***

Children of color may be less likely to experience high-quality classrooms, particularly in the domains of emotional and instructional support (Aguiar & Aguiar,

2020; Valentino, 2018). However, high-quality ECE may still serve as a buffering mechanism against the effects of racialized oppression for children of color and children who speak languages other than English at home as they prepare for formal schooling. For example, Bassok (2010) found that, controlling for income, Black children were more likely to benefit from high-quality early childhood education, compared to White children. Further, instructional quality was a stronger predictor of achievement for Black children than White children (Burchinal et al., 2011b). Empirically examining the profiles of classroom quality in Head Start may help to provide insight into how various aspects of quality pattern together and are experienced by children of different racial/ethnic groups.

Additionally, in Head Start, 30% of children are dual language learners, and the most common language spoken besides English is Spanish (Head Start Program Information Report, 2020). However, only 15% of Head Start non-supervisory staff are bilingual, meaning DLL children's access to a teacher that shares their language is limited. Of teachers who speak a language besides English, 78% are proficient in Spanish (Head Start Program Information Report, 2020). Some researchers have examined the relation between classroom DLL composition and classroom quality. Sawyer and colleagues (2016) found that lower Early Childhood Environmental Rating System (a measure of structural quality) scores were associated with higher proportions of DLLs, suggesting that children in classrooms with higher proportions of DLLs experienced lower classroom quality. However, no such relation was found between DLL

composition and CLASS scores. However, no studies to my knowledge have investigated DLL composition and classroom quality profiles.<sup>2</sup>

### ***Teacher-Child Match***

Teacher-child racial/ethnic match has largely been studied in the K-12 context but has been examined in some early childhood (child care, pre-k, Head Start) studies (Burchinal & Cryer, 2003; Choi & Dobbs-Oate, 2016; Downer et al., 2016; Ewing & Taylor, 2009; Graves & Howes, 2011; Ho et al., 2012; Howes et al., 2011; Howes & Shivers, 2006; Markowitz et al., 2020; Saft & Pianta, 2001). Generally, these studies have attempted to examine the impact of teacher-child racial/ethnic match on child outcomes, and about half have found significant findings (Burchinal & Cryer, 2003; Choi & Dobbs-Oates, 2016; Ewing & Taylor, 2009; Graves & Howes, 2011; Howes et al., 2011). Ho and colleagues (2012) found that White teachers' appeared to rate White children more favorably than non-White children. Howes & Shivers (2006) found that racial/ethnic mismatch was associated with conflictual interactions, and Saft & Pianta (2001) found that teachers rated their relationships with matched children more positively. These findings suggest that we still do not have a clear conceptualization of the implications of match. Additionally, more recently, researchers have begun to investigate the potential mediating and moderating mechanisms through which racial/ethnic match may benefit children. For example, Markowitz and colleagues (2020) found that match significantly predicted family engagement, which in turn reduced child

---

<sup>2</sup> It is important to note that these various aspects of identity are artificially isolated and in children's lives are always experienced together. The intersection of race/ethnicity and language is important to consider in early childhood and is a key future direction.

absences. Understanding racial/ethnic match in the context of quality profiles and child outcomes is another potential avenue for untangling the implications of match. Research on language match is still emerging, but studies by Downer and colleagues (2016) and Wright and colleagues (2017) suggest that language match may positively impact children's outcomes, over and above racial/ethnic match. Specifically, Downer and colleagues (2016) found that Latine DLLs with limited English proficiency who experienced racial/ethnic match grew more in their early literacy skills than those with White teachers.

### *Classroom Language Composition*

Classroom composition is generally thought of as an important feature of children's classroom experiences, particularly when it comes to access to resources, including quality (Aguiar & Aguiar, 2020). In this study, I specifically focus on DLL composition as a potential predictor of quality profiles, which has yet to be examined. Previous research examining DLL ECE classroom composition and the CLASS has primarily found null findings (Downer et al., 2012; Justice et al., 2008; Sawyer et al., 2016). For example, Justice and colleagues (2008) found that language classroom composition did not significantly predict high-quality instruction. Even studies using a contextual-specific measure of classroom quality (Sawyer et al., 2016) have found null relations between the DLL composition and classroom quality. However, White and colleagues (2020) found that both DLL-specific and general quality, as measured by the CLASS, positively predicted executive functioning among DLLs. Additionally, Downer and colleagues (2012) found that the CLASS functioned equivalently across various

levels of classroom DLL composition. Therefore, testing the relation between DLL composition in the context of quality profiles may provide additional insight into this relation.

### **Teachers as Socializers**

Teachers are key socializers and play a large role in ECE, therefore their perspectives are critical for children's experiences. Though the impetus of this study is to examine patterns of classroom quality and their relation to Head Start children's characteristics (dual language composition and school readiness), teachers play an oversized role in these constructs. Therefore teacher positionality (i.e., teacher's identities and the power structures that they intersect with) is a critical component of understanding classroom quality. The CLASS operationalizes classroom quality primarily through teacher behaviors, and teacher ratings of children's school readiness, like those used in this study, reflect a teacher's perception of that child's social skills and approaches to learning. Though teacher's ratings of children's skills are inherently inseparable from the teacher's perspective and therefore not objective, they are essential to prioritize because of the massive amounts of power teachers have over children in the early childhood context. For example, when completing a measure of children's social skills that are rooted primarily in White patterns of behavior and interaction, a White teacher may not fully capture all of the other aspects of that child's social skills. Therefore this rating is not culturally responsive and likely does not capture the child's true social skills. However, it likely does align with the teacher's perception of that child's social skills, particularly as related to the teacher's training in child development. In sum, teachers'

perceptions of children influence the way they interact with those children, their interactional quality shapes the daily flow of the classroom and children's experiences, and the quality of instruction influences children's school readiness.

### **School Readiness**

The term school readiness is often used as a catch-all to describe the skills that children need to be ready for the context of formal schooling. Generally, it contains both social and academic domains, but in recent years, the importance of executive function and physical health in academic and whole child functioning has become clear, and these domains have been added, as well (Aguiar & Aguiar, 2020). In this study, I focus specifically on two central aspects of school readiness, social skills and approaches to learning, because both of these readiness skills are central to children's functioning in the classroom. They are closely tied to children's competence at formal schooling, and the constructs themselves and teacher perspectives about children's skills in these areas may be impacted by bias (Brandmiller et al., 2020; Eggum-Wilkens et al., 2014; McDermott et al., 2018).

Historically, schools were developed to educate children that were usually White, wealthy, and male. As the system expanded into early childhood, it was primarily based on European models (i.e., Montessori, Kindergarten, etc.), and these roots in predominantly White, Western, and Protestant values persist in our education system (Prochner & Nawrotzki, 2019). Delpit (1988) describes the pervasiveness and persistence of the culture of power as a hidden curriculum that aligns with the values, skills, and interaction styles of those with power, i.e., White people in a White-dominant society.



Social skills and approaches to learning may be thought of as two manifestations of the culture of power in classrooms. It is important that children are explicitly taught this hidden curriculum and these skills in order to navigate classrooms and schools (Apple, 1971). Making these implicit skills explicit is one way of preparing children of color for the inequitable system they are about to enter, but does not negate the importance of making systems more culturally responsive or rebuilding them altogether for children of color to thrive.

To function well in a classroom, students need to learn how to interact with peers outside of their home environment, many for the first time. They also need to learn how to interact with adults beyond family members, listen to instructions from teachers, and focus on teacher-prescribed tasks. Yet, children and teachers both bring their whole selves to school including their beliefs, biases, and emotions. Unfortunately, children's expression of emotions through their behavior in the classroom can be misinterpreted, especially for Black children who are often adultified by teachers (Epstein et al., 2017). For example, Black girls, when high in social competence may be seen as self-sufficient and not in need of additional emotional support, when developmentally this would be an important need (Epstein et al., 2017). Or Black boys who demonstrate hyperactive behaviors may be viewed as troublemakers by teachers, rather than simply being energetic (Epstein et al., 2017). Further, skills and behaviors that are specific to the classroom context (i.e., raising of hands before speaking, taking turns speaking, directing eye contact at the teacher) may align more easily with the home culture of White children as behavioral signals of interest in learning; however, these cues may not be implicitly

understood by children from other racial/ethnic backgrounds with different patterns of communication. Moreover, these behaviors are not the only ways to express engagement with learning. Managing these social skills and approaches to learning are essential to the ways of being required of students to be successful in classrooms, but are rarely explicitly taught, exacerbating inequities created by hidden curriculum.

***Social Skills, Approaches to Learning, Relations to Quality***

Social skills are a broader category which encompasses social skills (interpersonal skills, self-control, verbal assertion) and problem behaviors (internalizing and externalizing problems; Gresham & Elliot, 1988). Quality of ECE has been shown to be associated with positive trajectories of children's social skills (Broekhuizen et al., 2016; Mashburn et al., 2008). Children's approaches to learning refers to the behavioral process by which children learn in the classroom. These are effortful learning behaviors, such as strategic planning, cooperation in group learning activities, or focus on learning activities (McDermott et al., 2011). They generally fall into three categories: competence motivation, attention/persistence, and attitudes toward learning (McDermott et al., 2002). In a statewide study of preschool children's learning behaviors, Domínguez and colleagues (2010) found that classroom organization positively predicted children's rate of learning behavior skill development over time. Additionally, in a nationally representative sample of Head Start children, classroom quality positively predicted approaches to learning which in turn predicted children's academic skills (Bustamante & Hindman, 2019). Further, in a study of Chinese children, Hu and colleagues (2017) found that emotional support positively predicted children's attention/persistence and learning

strategy. Therefore, it is important to understand the potential relations between patterns of quality and children's approaches to learning.

### **Present Study**

Viewing this literature through the lens of anti-racist early childhood education theory, and paying attention to the nexus of research, policy, and practice, there are key gaps. Specifically, it is essential to understand patterns of classroom quality in Head Start, examine the sociocultural context of the classroom, and test any subsequent associations with children's school readiness. The current study addresses gaps in this literature by: a) using a latent profile analysis to create profiles of classroom quality using the CLASS as the sole quality measure in profile creation with a nationally representative Head Start sample, b) examining the sociocultural context of the classroom in this analysis (i.e., language composition of the classroom, child race/ethnicity, teacher-child racial/ethnic or language match), and c) examining profiles as predictors of social skills and approaches to learning outcomes. Specifically, this study will address the following research questions:

**Research Question 1:** What patterns of classroom quality exist in a nationally representative Head Start sample? What are the teacher and child characteristics within these profiles?

**Research Question 2:** Does classroom DLL composition predict classroom quality profile membership?

**Research Question 3:** Do profiles of classroom quality and sociocultural context (i.e., child race/ethnicity, child language, teacher-child racial/ethnic match, teacher-child language match) predict children's social skills and approaches to learning?

## CHAPTER 7

### STUDY 2 METHOD

#### **Participants**

Using the Head Start Family and Child Experiences Survey (FACES) 2009 cohort, this study examined CLASS observations from 486 Head Start classrooms and assessments from 2906 children from the second year of the three-year longitudinal study. The study used a probability sampling method to collect a national Head Start sample and to select Head Start programs, centers, and classrooms and used equal-probability sampling at the child-level to select children in each classroom. Teachers in the study were 99.4% female and 39% reported speaking Spanish. The racial/ethnic makeup consisted of 35.7% White teachers, 28.9% Black teachers, 19.5% Latine teachers, 12.1% teachers who identified as “Another Race,” 2.0% Asian or Pacific Islander, and 1.8% American Indian/Alaska Native teachers. Classroom DLL composition was on average 31.5% ( $SD = .36$ ) and 921 children were Latine DLLs. Children were about half female (49.8%) and 26% spoke Spanish in the home. The racial/ethnic makeup consisted of 39.6% Latine, 31.8% Black children, 20.6% White children, 5.4% Multiracial, 1.7% Asian or Pacific Islander, 0.6% American Indian/Alaska Native (see Table 3 for teacher-child racial/ethnic match). On average, children were 45.77 months old ( $SD = 6.54$ ) at T1. Mother’s education level was also collected, and 36.4% had obtained less than a high school diploma, 34.2% graduated high school or

received a GED, 23.6% had attended some college, and 5.8% had obtained a bachelor's degree or higher level of education.

## **Procedure**

The FACES 2009 cohort involved four waves of data collection. The first wave was collected in the fall of 2009 (T1), the second was in the spring of 2010 (T2), and the third was spring 2011 (T3). There was also a fourth wave, in which children were followed from pre-k to kindergarten. The current study uses various measures from T1, T2, and T3. For survey data, teachers and parents filled out a questionnaire online regarding each participating child. A paper version of the questionnaire, that matched the online version as closely as possible, was also available for teachers or parents who requested a paper version or who did not have access to the internet. Classroom observations were conducted during wave two. Observations were minimally intrusive and involved observers conducting their ratings in the morning for at least four hours (for additional details, see Malone et al., 2013).

## **Measures**

### ***Social Skills***

Children's social skills were operationalized as cooperative classroom behavior and rated by teachers using items from the Social Skills Rating System (Gresham & Elliot, 1990) and the Personal Maturity Scale (Entwisle et al., 1987) at all three time points on a scale from 0 ("never") to 2 ("very often"). Social skills included behaviors such as helping put things away, following the rules of games, and complementing peers. A sum score composite including the same 12 items was created at each time point. The

scale demonstrated good reliability (Cronbach's alphas: .88-.90). A composite was formed using a sum score of 12 items.

### ***Approaches to Learning***

Children's approaches to learning was measured on a scale of 1 ("never") and 4 ("very often"), using the teacher-reported Approaches to Learning Scale created by the U.S. Department of Education. The scale captures motivation, attention, organization, persistence, and independence in learning (e.g., "pays attention well," "show eagerness to learn new things," "persists in completing tasks") and demonstrated good reliability (Cronbach's alpha: .89; Duncan et al., 2007). Scores on this scale have also been associated with the Preschool Learning Behavior Scale (McDermott et al., 2000). A composite was created using a mean score of five items.

### ***Demographics***

Child demographic data were obtained through the parent survey. For this study, parents provided their child's race/ethnicity, language, gender (0 = girls, 1 = boys), age, and mother's education. Children were identified as Multiracial if their parents indicated being of different race/ethnicities. Teacher demographic data were gathered through the teacher survey. For this study, teachers provided their race/ethnicity, and language(s) spoken, and gender. Teachers also reported on the number of DLLs (DLL children of any language background) in their classroom, which was then converted into a proportion of DLLs in a classroom by the Mathematica team who executed FACES 2009. Unfortunately, classroom racial/ethnic composition could not be examined due to limitations in sampling.

## **Analytic Plan**

### ***Preliminary Analyses***

Descriptive statistics and correlations among study variables were examined, including means, standard deviations, skewness, kurtosis, scatterplots, and bivariate correlations.

### ***Primary Analyses***

**RQ 1 Profiles of Classroom Quality.** To address research question 1, latent profile analysis (LPA) were used to identify the number of latent profiles that best capture the heterogeneity of classroom quality in the data. A series of models was tested and compared ranging from a 1-class model to an 8-class model. The best model was identified using model fit indices and theoretical interpretation of classes. Currently there is not a gold standard in the field for fit of LPA models, and therefore theory also plays an important role in class selection. Model fit indices included the Bayesian information criteria (BIC) for relative fit and the significance of the Lo-Mendell-Rubin (LMR) likelihood ratio test. Lower values for the BIC indicated better relative fit and a significant LMR indicates that the LPA class model being tested fits the data better than the one with fewer classes (Nylund et al., 2007). Further, entropy was examined to determine the separation or distance between classes, with values greater than .8 on a class of 0-1 suggesting greater distinction between classes (Masyn, 2013).

**RQ 2 Determining Likelihood of Quality Profile Membership.** To address research question three, multinomial logistic regression was conducted to determine the



likelihood of DLL classroom composition being associated with a particular quality profile.

**RQ 3 Relating Quality Profiles to School Readiness.** The connection between quality profiles and children's school readiness outcomes was examined using a hierarchical linear modeling approach to enter level 1 predictors (child race/ethnicity, child language, teacher-child racial/ethnic match, and teacher-child language match) and level 2 profiles. Analyses were conducted separately for children's social skills and approaches to learning to isolate the relation of quality to the social and academic components of school readiness.

## CHAPTER 8

### STUDY 2 RESULTS

#### *Preliminary Analyses*

Descriptive statistics indicated that variables were relatively normally distributed and correlated in the expected directions (see Tables 1 and 2). Chi-square analyses were conducted to examine potential issues of multicollinearity among the level-1 demographic predictors. Specifically, the relation between teacher-child racial/ethnic match for Latine, teacher-child language match (Spanish), child race/ethnicity (Latine), and child language (Spanish) were examined. The chi-square tests were significant across all these variables, indicating a high degree of association between these variables and limiting their ability to be used together as predictors in analyses. Therefore, to remain as close as possible to the research questions and aims of the study to understand quality and factors at the teacher level, the two match variables were maintained as predictors.

#### *Primary Analyses*

**RQ 1 Profiles of Classroom Quality.** To address RQ 1, I conducted latent profile analyses for 1-8 class solutions. However, after the 4-class solution, the log likelihood no longer replicated, despite increasing start values up to 1000. This indicated that even if the start values were increased further, the class-solution likely would not be high-quality. Therefore, I investigated the 1- through 4-class solutions (see Table 4 for classification criteria). The selection of a class solution is determined by considering the information criteria, measurement certainty, ratio tests, and figures for each class, and weighing each of these pieces of evidence in the context of theory. The information

criteria showed the largest drop in likelihood values and BIC between the 1- and 2- class solution and smaller changes in BIC between the 2- to 3- class solutions and the 3- to 4- class solutions. The measures of measurement certainty showed that all the class solutions had relatively high entropy and that the entropy increased with each class solution and the 4-class solution had the highest entropy at .909. The average posterior probabilities indicated that the classes were distinct, however the class sizes were only reasonable for the 2- and 3-class solutions. Finally, the LMR was not significant in all solutions and the BLRT was significant for all four class solutions.

The next step was to examine the mean-centered figures of the 2-, 3-, and 4-class solutions to highlight any potential patterns (see Figures 1-3). The 2-class solution showed one high-quality class with low negative climate and one low-quality class with higher negative climate. The 3-class solution profiles differed from one another quantitatively with a low, medium, and high-quality class, but no distinctive qualitative differences. However, in the 4-class solution, another class emerged. There were similar low quality, moderate quality, and high-quality classes, but there was also a class with high instructional quality and lower classroom organization and relatively low emotional support (Class 4 Figure 3). Though this 4-class solution is theoretically interesting, it was not empirically feasible with the .004% class size for class 3. Comparing the 2- and 3-class solutions, the solution that provided insight into the heterogeneity of classroom quality through profiles and had higher entropy (.887) was the 3-class solution.

Within the 3-class solution were three distinct classes, though the patterns of indicators were similar across classes (see Figure 4). The first class (profile 1 moderate

emotional support, low classroom organization, and higher negative climate), containing 23% of classrooms, was characterized by relatively low mean-centered scores across quality indicators and higher negative climate. Specifically, it contained classrooms with very low scores on the instructional support indicators (concept development, feedback, and language modeling), moderate scores on the emotional support domain with moderate positive climate and teacher sensitivity, but also relatively higher negative climate, and lower than average classroom organization indicators. The second class (profile 2 high emotional support, classroom organization, and average instructional support), also containing 23% of classrooms, was relatively high-quality with scores above the sample mean on emotional support (low score on negative climate) and classroom organization indicators and scores right on the sample mean on instructional support indicators. The third class (profile 3 high emotional support, classroom organization, and lower instructional support), containing 54% of the classrooms, was characterized by moderate quality, relatively high scores on emotional support (low score on negative climate) and classroom organization indicators and relatively lower scores on instructional support.

### **RQ1a Descriptives of Quality Profiles**

The next step was to examine the demographic characteristics of the children within the classroom quality profiles. This involved hard classifying the posterior probabilities and assigning each classroom to a latent class, then examining the characteristics of the teachers and children in that classroom (see Table 5). The majority of children across racial/ethnic groups, languages, and sex were in profile 3 (high

emotional support, classroom organization, and lower instructional support). However, there were nuances in the distribution of groups of children across the profile 1 (moderate emotional support, low classroom organization, and higher negative climate) and profile 2 (high emotional support, classroom organization, and average instructional support) (see Tables 5). For example, only 18% of Black and Multiracial children and 22% of Latine children were in profile 2, as compared to 29% of White children. On the other hand, 31% of Black and 23% of Multiracial children were in the profile 1, as compared to 15% of Latine children and 14% of White children. Additionally, 62% of Latine DLLs were in the profile 3. Children who shared the racial/ethnic background of their teacher and those who did not were similarly distributed, with roughly 57-58% in the profile 3 and 20-21% in the profile 1 and profile 2. Additionally, the majority of teachers across racial/ethnic groups were in the profile 3, however, 29% of White and 26% of Latine teachers were in profile 2, and only 12% of Black teachers. Additionally, 31% of Black teachers were in profile 1, as compared to 16% of Latine and 11% of White teachers.

Another way to conceptualize the distribution of teachers and children of various racial/ethnic groups was to examine their representation within the profiles relative to their representation in the sample (Table 6). Relative to their representation in the sample, Latine children were overrepresented in the profile 3 and profile 2 (by 6 and 2 points respectively). White children were overrepresented in profile 2 (by ~6 points) and underrepresented in profile 3 and profile 1 (by ~1 and 7 points respectively). Black children were underrepresented in profiles 3 and 2 (by ~2 and 3 points respectively) and overrepresented by 20 points in profile 1. Similarly, relative to their representation in the

sample, Latine teachers were overrepresented in profile 3 and profile 2 (by 4 and 7 points respectively). White teachers were overrepresented in profile 3 and profile 2 (by ~7 and 17 points respectively) and underrepresented in profile 1 by ~15 points. Black teachers were underrepresented in profile 3 and profile 2 (by ~3 and 10 points respectively) and overrepresented by 31 points in profile 1.

### **RQ 2 Determining Likelihood of Quality Profile Membership**

To address research question 3, I used the R3STEP procedure in Mplus to predict latent class membership from the classroom proportion of DLLs. The R3STEP procedure in Mplus uses a multinomial logistic regression approach that allows for the uncertainty around class membership to be maintained. Results showed that the proportion of DLLs in the classroom was not a significant predictor of latent class membership (profile 1 vs. profile 2  $\beta = .787, p = .077$ ; profile 1 vs. profile 3  $\beta = .700, p = .149$ ; profile 3 vs. profile 2  $\beta = .087, p = .814$ ).

### **RQ 3 Relating Quality Profiles to School Readiness**

Mplus currently does not allow a level 2 latent categorical variable, such as the one created in research question one, to be embedded in multilevel analyses. Therefore, to address this research question, I created three variables based on the posterior probabilities of profile assignment to represent the profiles as closely as possible. Then the posterior probabilities were assigned at the child-level in order to meet the current constraints of Mplus (i.e., each child had an estimate of being in a particular profile, based on the estimate assigned to their teacher in the profile analyses conducted in research question 1). This format facilitated the Mplus modeling and was the most

parsimonious version of the model given that partitioning the variance between individual and classroom level was not a central aim of the study. Next, these variables were entered into the model as predictors of children's school readiness outcomes in combination with other predictors through a series of regression models.

TYPE=COMPLEX was used to address the nested nature of the data, and an MLR estimator was used. There was multicollinearity among the predictors when child demographics (race/ethnicity and language) and teacher-child match on those characteristics were included in the same model. Therefore, child demographics and match as predictors of school readiness were examined in two separate models, while controlling for school readiness at time 1. Child gender was also included as a child demographic covariate due to its centrality to social skills and approaches to learning in young children (Graves et al., 2012; Li-Grining et al., 2010).

First, T1 social skills and the quality profile predictors were entered into the model, and the profile 1 (moderate emotional support, low classroom organization, and higher negative climate) significantly negatively predicted T2 social skills (see Table 7). Building on this model, child demographic predictors were added. Profile 1 significantly negatively predicted T2 social skills, and child gender was significantly related to T2 social skills, such that boys had lower social skills. Next, a model with match characteristics and profile predictors was estimated, and these were not significant predictors of social skills.

The same approach was applied to approaches to learning (see Table 7). First, the T1 control and quality profiles were entered into the model, and the profile 1 was

significantly negatively related to T2 approaches to learning. Next, a model with the child-level demographic predictors was added, and again profile 1 negatively predicted approaches to learning. Child gender was also a significant predictor of approaches to learning, such that boys had significantly lower approaches to learning than girls. Third, a model was estimated with the match variables as predictors, and profile 1 was negatively associated with approaches to learning.



## CHAPTER 9

### STUDY 2 DISCUSSION

#### **Summary of Findings**

In this study I examined patterns of classroom quality as measured by the CLASS and the relation between these profiles and child school readiness as well as classroom composition of dual language learners. The 3-class solution was supported both empirically and theoretically, suggesting that in a nationally representative Head Start sample there were three profiles. Profile 1 (23% of classrooms) was characterized by moderate emotional support, low classroom organization, and higher negative climate. Profile 2 (23% of classrooms) was characterized by high emotional support, classroom organization, and average instructional support. Profile 3 (54% of classrooms) was characterized by high emotional support, classroom organization, and lower instructional support. Though the majority of the sample was in profile 3, relative to other racial/ethnic groups there was a higher proportion of White teachers and children that were in profile 2. Additionally, there were higher proportions of both Black teachers and children in profile 1 relative to Latine and White teachers and children. Additionally, DLL classroom composition was not a significant predictor of quality profile membership. Finally, profile 1 significantly predicted children's social skills and approaches to learning.

#### **Explanation of Findings**

##### **RQ 1: Quality Profiles**

In this study, I found evidence of three classroom quality profiles. Previous studies using a pattern centered approach to examining classroom quality have found a range of profile compositions in terms of the patterns within the profiles and number of profiles. The profiles found in this study aligned with previously literature in that four previous studies found evidence of three profiles and over half of previous studies found evidence a low, moderate, and high-quality profiles in some form or fashion (Hoang et al., 2019; Iruka & Morgan, 2014; Lippard et al., 2019; Salminen et al., 2018). The profiles were created using individual indicators, allowing for variability of indicators within the content of these three profiles. Additionally, mean centering these indicators highlighted differences in indicator patterns. For example, within instructional support (IS; i.e., concept development, feedback, language modeling) in profile 1, language modeling was lower than the mean relative to the two other IS indicators. In profile 2, language modeling was the highest above the mean relative to the other IS indicators. For profile 3 (high emotional support, classroom organization, and lower instructional support) language modeling was roughly the same as concept development and feedback relative to the mean. This pattern could suggest that language modeling is an important component of IS, particularly because of the tie between children’s language development and social development (Clayton et al., 2022). Additionally, negative climate was relatively similar to the mean for profiles 2 and 3, but above the mean for profile 1. This combination of indicators may provide insight in the connection to outcomes.

### **RQ 1a: Demographics and Quality Profiles**

To understand the distribution of individuals across quality I examined the distributions of teacher and child race/ethnicity across profiles. In profile 2, White teachers and children were overrepresented, Latine teachers and children were overrepresented, and Black teachers and children were underrepresented. For profile 3 Latine teachers and children were overrepresented. However, for profile 1, White and Latine teachers and children were underrepresented, and Black teachers and children were overrepresented. This pattern of findings is not uncommon in observational measures in education (Gill et al., 2016). Multiple studies have found that teachers of students who are low-income, children of color, or have lower achievement ratings, have lower observational scores (Borman & Kimball, 2005; Chaplin et al., 2014; Whitehurst, Chingos, & Lindquist, 2014). However, it is difficult to determine the source of these patterns. There are multiple empirical and theoretical possibilities. First, observational measures introduce a human observer. Additionally, the measure could be capturing a biased conception of quality that does not capture other manifestations of teacher classroom quality. Better measurement is needed to assess these potential sources of bias.

The implications of the racial/ethnic makeup of these profiles is

### **RQ 2: DLL Composition & Quality Profiles**

Classroom composition of DLLs was not a significant predictor of quality profile membership. Previous literature on composition of DLLs and DLL status for individual children more broadly is mixed on this characteristic's statistical relation to quality or individual teachers (Downer et al., 2012). These findings could be for multiple reasons,

some tied to measurement and some substantive. First, the DLL composition variable in this study was a teacher rated variable and could include a DLL of any language. The overwhelming majority of DLLs in the US are Spanish speaking DLLs, who also tend to be Latine. Due to the intersection of language and race/ethnicity in the US there may be specific racialization that is not captured by a broad DLL composition item. Language match may have also played a role in this relation. Interestingly there were more Spanish speaking DLLs who had a language match with their teachers than Latine children with racial/ethnic match despite a large overlap in these populations.

### **RQ 3: Quality Profiles & Child School Readiness**

Next, I examined the quality profiles as predictors of school readiness and found that the profile 1 significantly negatively predicted children's social skills and approaches to learning. The finding in this study that the profile 1 was associated with lower social skills was relatively unexpected. Evidence of the connection between high quality classrooms (similar to profile 2) and better social skills has been found in both variable centered (i.e., Broekhuizen et al., 2016; Mashburn et al., 2008) and found in pattern centered analyses (i.e., Curby et al., 2009). For example, Curby and colleagues (2009) found that quality profiles with the highest emotional support showed greater social competence. Additionally, multiple meta-analyses have found that low quality (similar to profile 1) is generally not associated with child outcomes and high quality generally significantly positively predicts child outcomes, such as social skills (Burchinal et al., 2016; Zaslow et al., 2010). However, a recent study by Aguiar and colleagues (2019) found that lower classroom organization was associated with lower social skills. Future

research should consider additional variables such as dosage of quality and child attendance (Aguiar et al., 2019). Additionally, Brophy-Herb (2007) found that teacher's negative behaviors and poor climate were associated with their negative child ratings. Because profile 1 showed the highest negative climate ratings this could provide an additional explanation of the significant negative association with social skills and approaches to learning. For approaches to learning, Hu and colleagues (2017) also found that low classroom organization negatively predicted approaches to learning. Whereas Bustamante & Hindeman (2019) found quality positively predicted gains in approaches to learning and Dominguez and colleagues (2011) found null relations. A potential future avenue of research is to examine gains in scores rather than rates. For example, Bustamante & Hindeman (2020) found gains in approaches to learning to be a strength of Latine children in Head Start.

This finding could be for a host of different reasons. One reason could be broader difficulty in measuring interactional quality. Additional observation windows may be needed to more reliably capture interactional quality in a way that is stable and generalizable to the practices of the teacher over time (Praetorius et al., 2014). For example, previous researchers have found that teacher behaviors may only account for 10%-45% of variability kindergarten CLASS scores (Mantzicopoulos et al., 2018). Another possibility is the continued issues with the predictive validity of the CLASS (Burchinal et al., 2016). One would expect the CLASS to be significantly positively *and* negatively associated with child outcomes. However, this is rarely the case in the literature (Burchinal et al., 2016). The third possibility could be the specific school

readiness outcomes in this study. Other studies that took a pattern centered approach tended to find significant relations between profiles and academic outcomes as opposed to social and behavioral child outcomes. Additionally, the use of individual indicators means that only 4 out of 10 indicators are focused on socioemotional constructs. Future research is needed to tease apart these relations, specifically focusing on dosage, gains/trajectories of outcomes, and additional important explanatory variables.

### **Implications**

This study focused on teacher's classroom quality at the nexus of teacher child interactions and the findings may have broader implications for technical assistance and professional development. First, the finding of three classroom quality profiles highlights the strengths of teachers across these profiles, particularly on positive climate, behavior management, and productivity where teachers in each profile were rated above a 4 on a 1-7 point scale (moderate according to the CLASS publishers; La Paro et al., 2004). Next, these profiles also highlighted the consistent struggle with IS indicators across each class that is echoed in the literature (Burchinal et al., 2016). This study also showed important findings related to the demographic characteristics of the teachers and children who were assigned to these classes. In line with anti-racist ECE theory, which posits that White individuals as well as ways of being are privileged in ECE spaces, higher proportions of White children and teachers were in the high-quality class. Within the structure of white supremacy, racial/ethnic groups are racialized in specific ways to uphold white power and dominance (Omi & Winant, 1986). In this study the high proportion of White

children and teachers in the profile 2, the relatively high proportion of Latine children and teachers in profile 3, and the relatively high proportion of Black children and teachers in profile 1 reflect this broader racialized hierarchy.

Additionally, there may be unique implications for Latine and Black teachers and children due to the unique racialization of these groups. For example, the racial/ethnic category of Latine is heterogenous reflecting a wide range of experiences and histories, including large populations of people with Mexican and Central American ancestry in the southwestern U.S. to those with ties to the Caribbean and South American. This diversity may lead to differential racialization. The overrepresentation of Latine children and teachers in profile 3, may be associated with these racialization factors and should be examined in future research. Next, Black children and families face pervasive Anti-Black racism which may be either reflected in Black children's overrepresentation in the lower quality profile (profile 1) or the profile itself may be impacted by Anti-Black racism in that observer's implicit biases may lead them to perceive Black teachers as harsher. Further research is needed to disentangle this relation.

### **Practice**

Teacher training and technical assistance may focus on the strengths of Head Start teachers in how across the profiles they create a positive climate for their students and build on this positive climate in other areas such as emphasizing intentional training on feedback and concept development, which have been shown to be important for social skills (Perlman et al., 2016). Additionally, the findings of this study that the low

classroom quality profile significantly negatively predicted social skills and approaches to learning highlights a need to emphasize training and support for these programs and classrooms. A key component of which is the combination of moderate ES and CO indicators with the nearly the lowest possible scores in the IS domain. Next, a broader contribution of this study is the use of anti-racist ECE theory and CRP as applied to classroom quality and child school readiness. These indicate that equity and cultural responsiveness are key missing indicators of interactional quality and areas where teachers and directors, particularly White teachers and directors may need substantial training. This training may take the form of content training on the history and present-day impacts of white supremacy and racism, critical reflection, development of cultural humility and distress tolerance in preservice and TA contexts.

### **Policy**

The findings of this study have implications for policy makers that interact with Head Start at the federal, state, and local level. Particularly the finding that more White children and teachers find themselves in profile 2 should signal to policy makers that a more systematic investigation is needed to determine as of 2022, the number of children of various racial/ethnic groups that have access to various patterns of classroom quality. Additionally, the finding that a higher percentage of Black teachers were in profile 1 signals a key area for future research. Profile 1 included higher negative climate and additional research is needed to examine potential observer bias in their interpretations of the behavior of Black teachers. For example, a biased observer may perceive a Black teacher's reaction to a student as harsh, when it may be a valid and culturally appropriate



response. Additionally, it highlights a need to quantitatively give teachers of color, particularly Black teachers, credit for the culturally responsive interactional quality they are likely doing. When this is not included in quality, it by default benefits White teachers and is “culturally responsive” to White children. A colorblind approach benefits white supremacy. Additionally, Head Start policy makers should expand their conception of quality to include CRP and anti-racist and anti-bias approaches and update standards to have this reflected in measurement. This requires tangible financial support at the federal level for measurement development by and for teachers of color.

Another area of policy impact is the Head Start Designation Renewal System, which outlines the conditions under which a Head Start grantee must re compete for a grant or continue without competition for a period of five years. Current Head Start policy outlines two policies for the designation renewal system related to the CLASS. First, if a grantee falls below the first quality threshold (6 ES, 6 CO, 3 IS), they will receive Office of Head Start support to improve quality (Office of Head Start, 2022). The second specifies that if a grantee’s average CLASS scores fall below 5 for ES, 5 for CO, and 2.3 (with upcoming changes in 2025) they are required to re compete for their grant. Future research should examine the proportion of Head Start grantees that fall are required to re compete based on these standards and determine if there are demographic groups of directors, teachers, or composition of centers that are overrepresented in these categories. This would provide additional insight into the equity impact of this policy. At a federal level OHS should conduct a current audit of children’s access to high quality centers including child characteristics such as race/ethnicity and language.

## **Future Directions and Limitations**

The findings of this study provided evidence for three quality profiles in the Head Start FACES 2009 cohort, allowed for the examination of individual distributions across these profiles, indicated that the profile 1 significantly negatively predicted children's social skills and approaches to learning, and that DLL classroom composition was not significantly associated with quality profile membership. However, there are limits to these findings. First, the findings should be explored specifically for Asian, Pacific Islander, Indigenous, and Multiracial children. Second, there are important limits on the conclusions that can be drawn from the examination of child demographic characteristics at the child level. We could not conduct significance tests due to the nested nature of the data and therefore additional research is needed to clarify the extent to which the composition of these profiles are significantly different from one another. Next, as discussed previously the DLL classroom composition variable was teacher reported and included children across all potential other languages precluding me from examining the uniquely racialized experience of Latine DLL composition as a predictor of quality profile.

Next, there are likely additional variables that should be investigated in the future in relation to the implications of quality profiles for children's school readiness. For example, additional classroom characteristics (teacher education, teacher-child ratio, teacher depression/stress, etc.). Additionally, future research should investigate longitudinal child exposure to quality profiles as this may provide additional insight into

the dosage of patterns of quality needed to buffer negative outcomes or enhance positive school readiness skills. Another important future area of research in relation to quality is the estimation of profiles as predictors of growth on social skills and approaches to learning. These profiles may be a better predictor of growth rather than these outcomes within time. Finally, future research should focus on examining quality profiles that account for culturally relevant pedagogies and practices. Explicitly measuring equity indicators of quality may be more likely to produce meaningful findings, particularly as related to sociocultural context variables.

## CHAPTER 10

### GENERAL DISCUSSION

Head Start classroom quality is central to the experiences of both teachers and children. For teachers, their ratings on classroom quality measures, as rated by the CLASS, may be used to assess their teaching and determine funding as it is used to hold Head Start grantees accountable at the federal level. For children, the ways that their teachers interact with them is central to their experiences in the classroom and has implications for their development of school readiness skills. All of these processes take place within the sociocultural context of the classroom and are impacted by the broader structures of racism and white supremacy in US society. The aim of this dissertation was to examine Head Start classroom quality through the lens of anti-racist ECE theory, necessitating a critical approach to teacher and child identity, classroom quality, and school readiness. First, an empirical assessment of measurement invariance across teacher race/ethnicity had yet to be published, requiring researchers and policy makers to assume, rather than know empirically, that the CLASS was capturing the same constructs across teacher groups. In the first study I investigated measurement invariance across teacher race/ethnicity (Black, Latine, White) on the CLASS to determine if there was systematic measurement bias on race/ethnicity. Additionally, studies of profiles of classroom quality had yet to focus on the sociocultural context of the Head Start classroom, specifically for outcomes that are tied to the hidden curriculum of school readiness. In the second study I examined profiles of classroom quality, DLL classroom composition as a predictor of profile membership, and quality profiles and demographic

characteristics (i.e., child race/ethnicity and teacher-child racial/ethnic and language match) as predictors of child school readiness (i.e., social skills and approaches to learning).

These two studies approach different aspects of the intersection of interactional quality and teacher's and children's Head Start experiences. In study 1 I found evidence of partial scalar invariance, indicating that means may be compared across teacher race/ethnicity (Latine, Black, and White). These findings suggest that the CLASS, though it does not give teachers (particularly teachers of color) any credit for culturally responsive or equity focused practices, it meets the minimal standard of not perpetuating measurement bias based on teacher race/ethnicity. This finding should provide a starting point for equity-focused investigations of the CLASS measure rather than the end point, particularly in the context of the importance of this measure for the ECE workforce. Next, in study 2 I found evidence of three profiles of classroom quality, profile 1 (moderate emotional support, low classroom organization, and higher negative climate), profile 2 (high emotional support, classroom organization, and average instructional support), profile 3 (high emotional support, classroom organization, and lower instructional support). I then examined the profiles that teachers and children were most likely to be in and found that White teachers and children were overrepresented in the profile 2. This aligns with the notion that the CLASS may be tailored to the experiences of these specific groups. Finally, the profile 1 negatively predicted child social skills and approaches to learning, meaning that low quality patterns may be detrimental for the

skills children need to function in the classroom over and above any child demographic or match characteristics.

Together and in light of anti-racist ECE theory, these two studies suggest that the existing paradigm of defining classroom quality through the CLASS is not sufficient to capture both the experiences and outcomes of teachers and children of color. Though the measure will continue to be used to compare means across teachers and centers, this does not indicate that it is the most fair or equitable tool to assess classroom quality.

Additionally, it remains to be seen if the new iteration of the CLASS will address the content issue of equity and cultural responsiveness in classroom quality as evidence from study 2 suggests that the CLASS is culturally responsive to White teachers and children. Additionally, though partial scalar measurement invariance across teacher race/ethnicity is positive, it is a starting point and baseline. Particularly when paired with the finding that the majority of classrooms were in profile 3, yet Black teachers and children were still overrepresented in profile 1 and White teachers and children were over-represented in profile 2. This finding should be examined in other samples better suited to analytically capture the likelihood of teachers and children of various racial/ethnic groups to be in particular quality profiles.

Through the lens of anti-racist ECE theory, these dissertation findings suggest I find that the current paradigm of measuring classroom quality maintains the status-quo. Further research is needed to better understand the potential mediating or moderating mechanisms of this relation between profile 1 and children's classroom behavior. Additionally, they suggest further empirical scrutiny is needed regarding the positionality

of raters of high-stakes observational measures. Results from this dissertation inform future research, encourage more critical understandings of the CLASS by ECE policy and decision makers, and additional praxis is needed to prevent analysis paralysis. Praxis in the ECE paradigm requires a critique of current approaches to minimize harm and advocating for the reimaging of systems that are designed for equitable experiences; *education as the practice of freedom* (hooks, 1994).

### **Sociopolitical Historical Context**

As I am writing this conclusion during the last week of May 2022, two mass shootings have taken place in the last two weeks. The first was a mass shooting by a white supremacist who killed 10 Black people in a grocery store in upstate New York. This is a gut-wrenching and all-too common and predictable tragedy. The people who were lost are irreplaceable, leaving legacies, family members, and a community in grief. They were targeted because of their race. This shooting was white supremacy in action and highlights the very real life or death stakes of working toward social justice and against racism. Then, just a 10 days later on May 24<sup>th</sup>, another gunman committed a mass shooting at an elementary school in Texas killing 2 adults and 19 children. This is a heartbreaking and predictable tragedy in a predominantly Latine community. On May 25<sup>th</sup> is also the two-year anniversary of the police murders of George Floyd and Dion Johnson, whose unjustified deaths energized a national movement toward anti-racism and social justice.

These brutal murders, the overwhelming and cruel power of the gun lobby and the politicians they buy off, and the permanent pain caused by these incidents is part of the

context of this dissertation. It highlights, that some of our most important members of society, our teachers are not only being asked to educate children when they go to work, but to defend children with their lives. Children go to school and in addition to fire drills, they participate in traumatizing active shooter drills, because the adults in their country refuse to protect them. Quality in ECE is and children's school readiness are important, but ultimately rely on the physical safety of students and teachers – requiring addressing this uniquely American problem to truly begin the process of shaping what happens inside classrooms.

“Schools scared to death.

The truth is, one education

under desks,

Stooped low from bullets;

That plunge when we ask

Where our children

Shall live

& how

& if

It takes a monster to kill children. But to watch

monsters kill children again

and again and do nothing

isn't just insanity—it's

inhumanity.



The truth is, one nation under guns.

What might we be if only we tried.

What might we become if only we'd listen.”

- Amanda Gorman

## REFERENCES

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213.
- Administration for Children and Families (2020). *Head Start Designation Renewal System: A Rule by the Children and Families Administration on 8/28/2020*. <https://www.federalregister.gov/documents/2020/08/28/2020-17746/head-start-designation-renewal-system>
- Aguiar, A. L., & Aguiar, C. (2020). Classroom composition and quality in early childhood education: A systematic review. *Children and Youth Services Review*, *115*, 105086.
- Aguiar, A. L., Aguiar, C., Cadima, J., Correia, N., & Fialho, M. (2019). Classroom quality and children's social skills and problem behaviors: Dosage and disability status as moderators. *Early Childhood Research Quarterly*, *49*, 81-92.
- Apple, M. W. (1971). The hidden curriculum and the nature of conflict. *Interchange*, *2*(4), 27-40.
- Austin, Lea J. E. 2021. "Advancing Financial Equity for the ECE Workforce: Routes forward for Better Compensation." Presentation at Racial, Economic, and Social Justice for the Early Care and Education Workforce: Pre-, During, and Post-COVID-19 Convening, January 12–13.
- Austin, L.J.E., Edwards, B., Chávez, R., & Whitebook, M. (2019). Racial Wage Gaps in Early Education Employments. Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley. Retrieved from <https://csce.berkeley.edu/racial-wage-gaps-in-early-education-employment/>
- Barajas-Gonzalez, R.G. (2021). Early Care and Education Workforce Stress and Needs in a Restrictive, Anti-Immigrant Climate. Urban Institute. [https://www.urban.org/sites/default/files/publication/104620/early-care-and-education-workforce-stress-and-needs-in-a-restrictive-anti-immigrant-climate\\_0.pdf](https://www.urban.org/sites/default/files/publication/104620/early-care-and-education-workforce-stress-and-needs-in-a-restrictive-anti-immigrant-climate_0.pdf)
- Barnes-Najor, J. V., Thompson, N. L., Cameron, A. F., Smith, T. M., Calac Verdugo, M., Brown, P. L., & Sarche, M. C. (2021). Cultural and Practice Perspectives on the Classroom Assessment Scoring System: Voices From American Indian and Alaska Native Head Start Programs. *Journal of Research in Childhood Education*, *35*(1), 162-183.

- Bassok, D., & Galdo, E. (2016). Inequality in preschool quality? Community-level disparities in access to high-quality learning environments. *Early Education and Development, 27*(1), 128-144.
- Bayly, B. L., Bierman, K. L., & Jacobson, L. (2021). Teacher, Center, and Neighborhood Characteristics Associated with Variations in Preschool Quality in Childcare Centers. In *Child & Youth Care Forum* (pp. 1-25). Springer US.
- Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., ... & Gill, S. (2008a). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development, 79*(6), 1802-1817.
- Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008b). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology, 20*(3), 821-843.
- Bihler, L. M., Agache, A., Schneller, K., Willard, J. A., & Leyendecker, B. (2018). Expressive morphological skills of dual language learning and monolingual German children: Exploring links to duration of preschool attendance, classroom quality, and classroom composition. *Frontiers in Psychology, 9*, 888.
- Bowlby, J., & Ainsworth, M. (2013). The origins of attachment theory. *Attachment theory: Social, Developmental, and Clinical Perspectives, 45*, 759-775.
- Brandmiller, C., Dumont, H., & Becker, M. (2020). Teacher perceptions of learning motivation and classroom behavior: The role of student characteristics. *Contemporary Educational Psychology, 63*, 101893.
- Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., ... & Lancet Early Childhood Development Series Steering Committee. (2017). Nurturing care: promoting early childhood development. *The Lancet, 389*(10064), 91-102.
- Broekhuizen, M. L., Mokrova, I. L., Burchinal, M. R., Garrett-Peters, P. T., & Family Life Project Key Investigators. (2016). Classroom quality at pre-kindergarten and kindergarten and children's social skills and behavior problems. *Early Childhood Research Quarterly, 36*, 212-222.
- Brophy-Herb, H. E., Lee, R. E., Nievar, M. A., & Stollak, G. (2007). Preschoolers' social competence: Relations to family characteristics, teacher behaviors and classroom climate. *Journal of Applied Developmental Psychology, 28*(2), 134-148.

- Brown, S., Souto-Manning, M., & Tropp Laman, T. (2010). Seeing the strange in the familiar: Unpacking racialized practices in early childhood settings. *Race Ethnicity and Education, 13*(4), 513-532.
- Bryan, N. (2017). White Teachers' Role in Sustaining the School-to-Prison Pipeline: Recommendations for Teacher Education. *Urban Rev 49*, 326–345.  
<https://doi.org/10.1007/s11256-017-0403-3>
- Bryant, D. M., Burchinal, M., Lau, L. B., & Sparling, J. J. (1994). Family and classroom correlates of Head Start children's developmental outcomes. *Early Childhood Research Quarterly, 9*(3-4), 289-309.
- 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.  
<https://doi.org/10.1016/j.ecresq.2012.02.001>
- Burchinal, M. R., & Cryer, D. (2003). Diversity, child care quality, and developmental outcomes. *Early Childhood Research Quarterly, 18*(4), 401-426.
- Burchinal, M., Kainz, K., & Cai, Y. (2011a). 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). Paul H. Brookes Publishing Co..
- Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S. L., McLoyd, V., ... & NICHD Early Child Care Research Network. (2011b). Examining the Black–White achievement gap among low-income children using the NICHD study of early child care and youth development. *Child Development, 82*(5), 1404-1420.
- Burchinal, M., Zaslow, M., & Tarullo, L. (Eds.). (2016). *Quality thresholds, features, and dosage in early care and education: secondary data analyses of child outcomes* (Vol. 81). Wiley.
- Bustamante, A. S., & Hindman, A. H. (2019). Classroom quality and academic school readiness outcomes in Head Start: The indirect effect of approaches to learning. *Early Education and Development, 30*(1), 19-35.

- Bustamante, A. S., & Hindman, A. H. (2020). Construyendo en la Fuerza: Approaches to learning and school readiness gains in Latino children served by head start. *Early Childhood Research Quarterly*, 52, 124-137.
- Chaplin, D., Gill, B., Thompkins, A., & Miller, H. (2014). *Professional practice, student surveys, and value-added: Multiple measures of teacher effectiveness in the Pittsburgh public schools*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, Regional Educational Laboratory MidAtlantic. <http://eric.ed.gov/?id=ED545232>
- Chávez-Moreno, L. C. (2021). Dual Language as White Property: Examining a Secondary Bilingual-Education Program and Latine Equity. *American Educational Research Journal*, 58(6), 1107-1141.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling: a multidisciplinary journal*, 14(3), 464-504.
- Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, 49(8), 566-578.
- Choi, J. Y., & Dobbs-Oates, J. (2016). Teacher-child relationships: Contribution of teacher and child characteristics. *Journal of Research in Childhood Education*, 30(1), 15-28.
- Clayton, R. J., Hein, S., Keller-Margulis, M. A., & Gonzalez, J. E. (2022). Associations between social skills, inattention, and English vocabulary skills of preschool Latinx dual language learners. *Journal of Research in Childhood Education*, 36(2), 219-238.
- Cumming, T. (2017). Early childhood educators' well-being: An updated review of the literature. *Early Childhood Education Journal*, 45(5), 583-593.
- Crane, J., Mincic, M. S., & Winsler, A. (2011). Parent-teacher agreement and reliability on the Devereux Early Childhood Assessment (DECA) in English and Spanish for ethnically diverse children living in poverty. *Early Education & Development*, 22(3), 520-547.
- Curby, T. W., LoCasale-Crouch, J., Konold, T. R., Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2009). The relations of observed pre-K classroom quality profiles to children's achievement and social

competence. *Early Education and Development*, 20(2), 346–372.  
<https://doi.org/10.1080/10409280802581284>

- Delaney, K. K., & Krepps, K. (2021). Exploring Head Start teacher and leader perceptions of the Pre-K Classroom Assessment Scoring System as a part of the Head Start Designation Renewal System. *Early Childhood Research Quarterly*, 55, 214-229.
- Delpit, L. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review*, 58(3), 280-299.
- Delpit, L. (2006). *Other people's children: Cultural conflict in the classroom*. The New Press.
- Dennis, R. M. (1995). Social Darwinism, scientific racism, and the metaphysics of race. *Journal of Negro Education*, 243-252.
- Domínguez, X., Vitiello, V. E., Fuccillo, J. M., Greenfield, D. B., & Bulotsky-Shearer, R. J. (2011). The role of context in preschool learning: A multilevel examination of the contribution of context-specific problem behaviors and classroom process quality to low-income children's approaches to learning. *Journal of School Psychology*, 49(2), 175-195.
- Downer, J. T., Goble, P., Myers, S. S., & Pianta, R. C. (2016). Teacher-child racial/ethnic match within pre-kindergarten classrooms and children's early school adjustment. *Early Childhood Research Quarterly*, 37, 26–38.  
<https://doi.org/10.1016/j.ecresq.2016.02.007>
- Downer, J. T., López, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: Applicability of the Classroom Assessment Scoring System in diverse settings. *Early Childhood Research Quarterly*, 27(1), 21-32.
- Education Week. (October, 2021) *Map: Where critical race theory is under attack*.  
<https://www.edweek.org/policy-politics/map-where-critical-race-theory-is-under-attack/2021/06>
- Eisenberg, N., Sadovsky, A., & Spinrad, T. L. (2005). Associations of emotion-related regulation with language skills, emotion knowledge, and academic outcomes. *New Directions for Child and Adolescent Development*, 2005, 109 –118.  
<http://dx.doi.org/10.1002/cd.143>

- Epstein, R., Blake, J., & González, T. (2017). *Girlhood interrupted: the erasure of black girls' childhood*. Center on Poverty & Inequality at Georgetown.
- Ewing, A. R., & Taylor, A. R. (2009). The role of child gender and ethnicity in teacher–child relationship quality and children's behavioral adjustment in preschool. *Early Childhood Research Quarterly, 24*(1), 92-105.
- Farewell, C. V., Quinlan, J., Melnick, E., Powers, J., & Puma, J. (2021). Job Demands and Resources Experienced by the Early Childhood Education Workforce Serving High-Need Populations. *Early Childhood Education Journal, 1*-10.
- Feagin, J. R., & Van Ausdale, D. (2001). *The first R: How children learn race and racism*. Rowman & Littlefield Publishers.
- Fernyhough, C. (2008). Getting Vygotskian about theory of mind: Mediation, dialogue, and the development of social understanding. *Developmental Review, 28*(2), 225-262.
- Gillborn, D., Warmington, P., & Demack, S. (2018). QuantCrit: education, policy, 'Big Data' and principles for a critical race theory of statistics. *Race Ethnicity and Education, 21*(2), 158-179.
- Goble, P., Eggum-Wilkens, N. D., Bryce, C. I., Foster, S. A., Hanish, L. D., Martin, C. L., & Fabes, R. A. (2017). The transition from preschool to first grade: A transactional model of development. *Journal of Applied Developmental Psychology, 49*, 55-67.
- Grant, A. A., Jeon, L., & Buettner, C. K. (2019). Relating early childhood teachers' working conditions and well-being to their turnover intentions. *Educational Psychology, 39*(3), 294-312.
- Graves Jr, S. L., & Howes, C. (2011). Ethnic differences in social-emotional development in preschool: The impact of teacher child relationships and classroom quality. *School Psychology Quarterly, 26*(3), 202.
- Graves Jr, S. L., Blake, J., & Kim, E. S. (2012). Differences in parent and teacher ratings of preschool problem behavior in a national sample: The significance of gender and SES. *Journal of Early Intervention, 34*(3), 151-165.
- Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology, 46*(5), 1062.

- Grooms, A. A., Mahatmya, D., & Johnson, E. T. (2021). The retention of educators of color amidst institutionalized racism. *Educational Policy*, 35(2), 180-212.
- Gill, B., Shoji, M., Coen, T., & Place, K. (2016). The content, predictive power, and potential bias in five widely used teacher observation instruments. *Mathematica Policy Research*.
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., ... & Hamagami, A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, 113(4), 461-487.
- Hardy, E, Joshi, P., Geronimo, K., Huber, R., Acevedo-Garcia, D. (2020). *Unequal Availability of Head Start: How Neighborhood Matters*. diversitydatakids.org.
- Harrison, L., Hurd, E., & Brinegar, K. (2021). But is it really about critical race theory?: The attack on teaching about systemic racism and why we must care. *The Middle School Journal*, 52(4), 2-3.
- Hauser-Cram, P., Sirin, S. R., & Stipek, D. (2003). When teachers' and parents' values differ: Teachers' ratings of academic competence in children from low-income families. *Journal of Educational Psychology*, 95(4), 813.
- Head Start Early Childhood Learning & Knowledge Center (2021). *Head Start Program Facts: Fiscal Year 2019*. Administration for Children & Families.
- Ho, H., Gol-Guven, M., & Bagnato, S. J. (2012). Classroom observations of teacher–child relationships among racially symmetrical and racially asymmetrical teacher–child dyads. *European Early Childhood Education Research Journal*, 20(3), 329-349.
- Howes, C., Guerra, A. W., Fuligni, A., Zucker, E., Lee, L., Obregon, N. B., & Spivak, A. (2011). Classroom dimensions predict early peer interaction when children are diverse in ethnicity, race, and home language. *Early Childhood Research Quarterly*, 26(4), 399-408.
- Howes, C., Fuligni, A. S., Hong, S. S., Huang, Y. D., & Lara-Cinisomo, S. (2013). The Preschool instructional context and child-teacher relationships. *Early Education and Development*, 24(3), 273–291. <https://doi.org/10.1080/10409289.2011.649664>
- Howes, C., & Shivers, E. M. (2006). New child–caregiver attachment relationships: Entering childcare when the caregiver is and is not an ethnic match. *Social Development*, 15(4), 574-590.



- Hu, B. Y., Dieker, L., Yang, Y., & Yang, N. (2016). The quality of classroom experiences in Chinese kindergarten classrooms across settings and learning activities: Implications for teacher preparation. *Teaching and Teacher Education*, *57*, 39-50.
- Hu, B. Y., Teo, T., Nie, Y., & Wu, Z. (2017). Classroom quality and Chinese preschool children's approaches to learning. *Learning and Individual Differences*, *54*, 51-59.
- Hu, L. T., & Bentler, P. M. (1998). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.
- Iruka, I. U., & Morgan, J. (2014). Patterns of quality experienced by African American children in early education programs: Predictors and links to children's preschool and kindergarten academic outcomes. *Journal of Negro Education*, *83*(3), 235–255. <https://doi.org/10.7709/jnegroeducation.83.3.0235>
- Jackson, J., Weidman, N. M., & Rubin, G. (2005). The origins of scientific racism. *The Journal of Blacks in Higher Education*, *50*, 66-79.
- Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51-68.
- Johnson-Staub, C. (2017). Equity Starts Early: Addressing Racial Inequities in Child Care and Early Education Policy. *Center for Law and Social Policy, Inc.(CLASP)*.
- Jones, N., Marks, R., Ramires, R., Rios-Vargas, M. (August, 2021). 2020 Census Illuminates Racial and Ethnic Composition of the Country. United States Census Bureau.
- Kachmar, S. P. (2008). Early childhood classroom quality and preschool learning behaviors.
- Kendi, I. X. (2019). *How to be an antiracist*. One world.
- Keum, B. T., Miller, M. J., Lee, M., & Chen, G. A. (2018). Color-Blind Racial Attitudes Scale for Asian Americans: Testing the factor structure and measurement invariance across generational status. *Asian American Journal of Psychology*, *9*(2), 149.

- Knight, G. P., Roosa, M. W., & Umaña-Taylor, A. J. (2009). *Studying ethnic minority and economically disadvantaged populations: Methodological challenges and best practices*. American Psychological Association.  
<https://doi.org/10.1037/11887-000>
- Kwon, K. A., Jeon, S., Jeon, L., & Castle, S. (2019). The role of teachers' depressive symptoms in classroom quality and child developmental outcomes in Early Head Start programs. *Learning and Individual Differences, 74*(June), 101748.  
<https://doi.org/10.1016/j.lindif.2019.06.002>
- La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal, 104*(5), 409-426.
- Lee, S. Y., Kim, R., Rodgers, J., & Subramanian, S. V. (2021). Treatment effect heterogeneity in the head start impact study: A systematic review of study characteristics and findings. *SSM-population health, 16*, 100916.
- Li, H., Liu, J., & Hunter, C. V. (2019). A meta-analysis of the factor structure of the classroom assessment scoring system (CLASS). *The Journal of Experimental Education, 88*(2), 265-287.
- Lifter, K., Foster-Sanda, S., Arzamarski, C., Briesch, J., & McClure, E. (2011). Overview of play: Its uses and importance in early intervention/early childhood special education. *Infants & Young Children, 24*(3), 225-245.
- Lippard, C. N., Choi, J. Y., & Walter, M. C. (2019). Profiles of classroom activity settings associated with Head Start children's receptive vocabulary. *Journal of Applied Developmental Psychology, 60*(November 2018), 65-75.  
<https://doi.org/10.1016/j.appdev.2018.11.002>
- LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly, 22*(1), 3-17. <https://doi.org/10.1016/j.ecresq.2006.05.001>
- Malone, L., Carlson, B. L., Aikens, N., Moiduddin, E., Klein, E., West, J., Kelly, A., ... & Rall, K. (2013). *Head Start Family and Child Experiences Survey: 2009 User's Manual*. Mathematica Policy Research.
- Mantzicopoulos, P., French, B. F., Patrick, H., Watson, J. S., & Ahn, I. (2018). The stability of kindergarten teachers' effectiveness: A generalizability study comparing the framework for teaching and the classroom assessment scoring system. *Educational Assessment, 23*(1), 24-46.

- Martinez-Cola, M. (2020). Visibly Invisible: TribalCrit and Native American Segregated Schooling. *Sociology of Race and Ethnicity*, 6(4), 468-482.
- Markowitz, A. J., Bassok, D., & Grissom, J. A. (2020). Teacher-child racial/ethnic match and parental engagement with Head Start. *American Educational Research Journal*, 57(5), 2132-2174.
- 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.
- Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. *The Oxford handbook of quantitative methods*, 2, 551.
- McLean, C., Austin, L.J.E., Whitebook, M., & Olson, K.L. (2021). Early Childhood Workforce Index – 2020. Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley. Retrieved from <https://cscce.berkeley.edu/workforce-index-2020/report-pdf/>
- McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (2000). Preschool Learning Behavior Scale. *School Psychology Quarterly*.
- McClelland, M. M., Tominey, S. L., Schmitt, S. A., & Duncan, R. (2017). SEL interventions in early childhood. *The Future of Children*, 33-47.
- Morgan, I., & Amerikaner, A. (2018). *Funding Gaps 2018: An Analysis of School Funding Equity across the US and within Each State*. Education Trust.
- Morris, P. A., Connors, M., Friedman-Krauss, A., McCoy, D. C., Weiland, C., Feller, A., ... & Yoshikawa, H. (2018). New findings on impact variation from the Head Start Impact Study: Informing the scale-up of early childhood programs. *AERA Open*, 4(2), 2332858418769287.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling: A multidisciplinary journal*, 14(4), 535-569.
- Office of Head Start (2020). *Use of Classroom Assessment Scoring System (CLASS®) in Head Start*. <https://eclkc.ohs.acf.hhs.gov/designation-renewal-system/article/use-classroom-assessment-scoring-system-class-head-start>

- Onwuachi-Willig, A., & Willig-Onwuachi, J. (2009). A house divided: The invisibility of the multiracial family. *Harv. CR-CLL Rev.*, 44, 231.
- Pages, R., Lukes, D. J., Bailey, D. H., & Duncan, G. J. (2020). Elusive longer-run impacts of head start: Replications within and across cohorts. *Educational Evaluation and Policy Analysis*, 42(4), 471-492.
- Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. *Child Development*, 72(5), 1534-1553.
- Perlman, M., Falenchuk, O., Fletcher, B., McMullen, E., Beyene, J., & Shah, P. S. (2016). A systematic review and meta-analysis of a measure of staff/child interaction quality (the Classroom Assessment Scoring System) in early childhood education and care settings and child outcomes. In *PLoS ONE* (Vol. 11, Issue 12). <https://doi.org/10.1371/journal.pone.0167660>
- Praetorius, A. K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2-12.
- Prochner, L., & Nawrotzki, K. (2019). The Origins of the Current Era of Early Childhood Care and Education. *The Wiley Handbook of Early Childhood Care and Education*, 7.
- Quality Compendium (2022). A Catalog and Comparison of Quality Improvement Systems (QIS). <https://qualitycompendium.org/>
- Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with Grade 2 achievement. *Early Childhood Research Quarterly*, 28(2), 437-449.
- Ricciardi, C., Manfra, L., Hartman, S., Bleiker, C., Dineheart, L., & Winsler, A. (2021). School readiness skills at age four predict academic achievement through 5th grade. *Early Childhood Research Quarterly*, 57, 110-120.
- Saft, E. W., & Pianta, R. C. (2001). Teachers' perceptions of their relationships with students: Effects of child age, gender, and ethnicity of teachers and children. *School Psychology Quarterly*, 16(2), 125.
- Salminen, J., Pakarinen, E., Poikkeus, A. M., & Lerkkanen, M. K. (2018). Development of pre-academic skills and motivation in kindergarten: a subgroup analysis

- between classroom quality profiles. *Research Papers in Education*, 33(4), 451–479. <https://doi.org/10.1080/02671522.2017.1353673>
- Sanders, C. R. (2016). *A chance for change: Head Start and Mississippi's Black freedom struggle*. UNC Press Books.
- Sawyer, B. E., Hammer, C. S., Cycyk, L. M., López, L., Blair, C., Sandilos, L., & Komaroff, E. (2016). Preschool teachers' language and literacy practices with dual language learners. *Bilingual Research Journal*, 39(1), 35-49.
- Schilder, D., & Sandstrom, H. (2021). *Racial, Economic, and Social Justice for the Early Care and Education Workforce: Pre-, During, and Post-COVID-19*. The Urban Institute.
- Schmit, S. & Walker, C. (February, 2016). Disparate Access: Head Start and CCDBG Data by Race and Ethnicity. *Center for Law and Social Policy, Inc.(CLASP)*. <https://www.ccf.ny.gov/files/1914/5625/2696/DisparateAccess.pdf>
- Seuring, J., Rjosk, C., & Stanat, P. (2021). Ethnic Classroom Composition and Minority Language Use among Classmates: Do Peers Matter for Students' Language Achievement?. *European Sociological Review*, 36(6), 920-936.
- Styck, K. M., Anthony, C. J., Sandilos, L. E., & DiPerna, J. C. (2021). Examining Rater Effects on the Classroom Assessment Scoring System. *Child Development*, 92(3), 976-993.
- Thompson, B., & Winsler, A. (2018). Parent–teacher agreement on social skills and behavior problems among ethnically diverse preschoolers with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 48(9), 3163-3175.
- Thorpe, K., Rankin, P., Beatton, T., Houen, S., Sandi, M., Siraj, I., & Staton, S. (2020). The when and what of measuring ECE quality: Analysis of variation in the Classroom Assessment Scoring System (CLASS) across the ECE day. *Early Childhood Research Quarterly*, 53, 274-286.
- Totenhagen, C. J., Hawkins, S. A., Casper, D. M., Bosch, L. A., Hawkey, K. R., & Borden, L. M. (2016). Retaining early childhood education workers: A review of the empirical literature. *Journal of Research in Childhood Education*, 30(4), 585-599.
- U.S. Department of Commerce, Census Bureau, American Community Survey (ACS), 2019. See *Digest of Education Statistics 2020*, table [202.20](#).

- Utt, J., & Tochluk, S. (2020). White teacher, know thyself: Improving anti-racist praxis through racial identity development. *Urban Education, 55*(1), 125-152.
- Valentino, R. (2018). Will public pre-K really close achievement gaps? Gaps in prekindergarten quality between students and across states. *American Educational Research Journal, 55*(1), 79-116.
- Vandenbroucke, L., Spilt, J., Verschueren, K., Piccinin, C., & Baeyens, D. (2018). The classroom as a developmental context for cognitive development: A meta-analysis on the importance of teacher–student interactions for children’s executive functions. *Review of Educational Research, 88*(1), 125-164.
- Venetsanou, F., & Kambas, A. (2010). Environmental factors affecting preschoolers’ motor development. *Early Childhood Education Journal, 37*(4), 319-327.
- Verhulst, F. C., & Akkerhuis, G. W. (1989). Agreement Between Parents' and Teachers' Ratings of Behavioral/Emotional Problems of Children aged 4–12. *Journal of Child Psychology and Psychiatry, 30*(1), 123-136.
- Watts, T. W., Nguyen, T., Carr, R. C., Vernon-Feagans, L., & Blair, C. (2021). Examining the Effects of Changes in Classroom Quality on Within-Child Changes in Achievement and Behavioral Outcomes. *Child Development, 92*(4), e439-e456.
- White, L. J., Fernandez, V. A., & Greenfield, D. B. (2020). Assessing classroom quality for Latino dual language learners in Head Start: DLL-specific and general teacher-child interaction perspectives. *Early Education and Development, 31*(4), 599-627.
- Whitehurst, G. J., Chingos, M. M., & Lindquist, J. M. (2014). *Evaluating teachers with classroom observations, lessons learned in four districts*. Washington, DC: Brown Center on Education Policy at the Brookings Institution.  
<http://eric.ed.gov/?id=ED553815>
- Wing, J. Y. (2007). Beyond black and white: The model minority myth and the invisibility of Asian American students. *The Urban Review, 39*(4), 455-487.
- Winston, A. S. (2020). Scientific racism and North American psychology. In *Oxford Research Encyclopedia of Psychology*.
- Wright, A., Gottfried, M. A., & Le, V. N. (2017). A kindergarten teacher like me: The role of student-teacher race in social-emotional development. *American Educational Research Journal, 54*(1), 78S-101S.

- Yip, Tiffany. 2021. "Addressing Inequities in Education: Considerations for Asian American Children and Youth in the Era of COVID-19." Society for Research in Child Development.
- Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Daneri, P., Green, K., ... & Martinez-Beck, I. (2016). I. Quality thresholds, features, and dosage in early care and education: introduction and literature review. *Monographs of the Society for Research in Child Development*, 81(2), 7-26.
- Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature. OPRE Report 2011-5. *Administration for Children & Families*.
- Zigler, E., & Styfco, S. J. (2010). *The hidden history of Head Start*. Oxford University Press.
- ZipRecruiter (January, 2022). What Is the Average Head Start Teacher Salary by State. <https://www.ziprecruiter.com/Salaries/What-Is-the-Average-Head-Start-Teacher-Salary-by-State>

APPENDIX A  
STUDY 1 & 2 TABLES



## STUDY 1 TABLES

**Table 1.**

*Study 1 Correlations*

	1	2	3	4	5	6	7	8	9	10	11	12
1. Concept Development	—											
2. Quality Feedback	0.759	—										
3. Language Modeling	0.609	0.754	—									
4. Positive Climate	0.448	0.448	0.371	—								
5. Negative Climate	0.210	0.278	0.349	0.425	—							
6. Teacher Sensitivity	0.396	0.438	0.515	0.620	-0.406	—						
7. Regard for Student Perspectives	0.338	0.316	0.450	0.489	-0.382	0.745	—					
8. Behavior Management	0.396	0.384	0.410	0.595	-0.461	0.606	0.475	—				
9. Productivity	0.381	0.337	0.402	0.537	-0.391	0.556	0.536	0.574	—			
10. Learning Format	0.363	0.352	0.493	0.461	-0.201	0.633	0.578	0.514	0.538	—		
11. Instructional Support	0.864	0.930	0.895	0.467	-0.317	0.506	0.416	0.443	0.417	0.455	—	
12. Emotional Support	0.446	0.468	0.530	0.806	-0.637	0.889	0.841	0.669	0.638	0.609	0.540	—
13. Classroom Organization	0.455	0.428	0.523	0.633	-0.414	0.717	0.638	0.820	0.849	0.831	0.526	0.764

*Note.* All correlations were significant  $p < .001$

**Table 2.***Study I Descriptive Statistics*

	Conce pt Devel opme nt	Qual ity Feed back	Langu age Mod eling	Positi ve Clima te	Negat ive Clima te	Teach er Sensit ivity	Regard for Student Perspec tive	Behav ior Mana geme nt	Prod uctiv ity	Lear ning For mat	Instr uctio nal Sup port	Emo tiona l Sup port	Classro om Organiza tion
<b>Valid</b>	408	408	408	408	408	408	408	408	408	408	408	408	408
<b>Missing</b>	91	91	91	91	91	91	91	91	91	91	91	91	91
<b>Mean</b>	2.072	2.268	2.454	5.316	1.241	4.635	4.489	5.041	4.925	4.009	2.265	5.300	4.658
<b>Std. Deviation</b>	0.678	0.721	0.847	0.669	0.455	0.667	0.678	0.745	0.838	0.856	0.672	0.498	0.678
<b>Variance</b>	0.460	0.520	0.718	0.448	0.207	0.445	0.460	0.554	0.701	0.733	0.452	0.248	0.459
<b>Skewness</b>	0.628	0.899	0.781	0.094	3.950	-0.290	-0.488	-0.567	0.403	0.144	0.792	0.892	-0.343
<b>Std. Error of Skewness</b>	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121
<b>Kurtosis</b>	0.172	1.148	0.643	0.504	26.799	-0.128	0.393	0.129	0.179	0.535	0.745	2.192	-0.053
<b>Std. Error of Kurtosis</b>	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241
<b>Minimum</b>	1.000	1.000	1.000	2.330	1.000	2.670	2.000	2.500	2.000	1.750	1.000	2.500	2.330
<b>Maximum</b>	4.250	5.000	5.000	7.000	5.670	6.330	6.250	6.750	7.000	6.000	4.560	6.380	6.170

**Table 4**

*CLASS Standardized Factor Loadings Across Measurement Invariance*

Invariance Step	Teacher Race/Ethnicity	Emotional Support Factor				Classroom Organization Factor			Instructional Support Factor			Covariances Among Factors			
		Positive Climate	Negative Climate	Teacher Sensitivity	Student Perspective	Behavior Management	Productivity	Learning Formats	Concept Development	Quality Feedback	Language Modeling	CO with ES	IS with ES	IS with CO	
105	Configural	Latine	0.834	-0.544	0.927	0.838	0.774	0.789	0.723	0.851	0.912	0.673	0.905	0.623	0.788
		Black	0.712	-0.642	0.839	0.772	0.748	0.795	0.645	0.824	0.900	0.753	0.914	0.632	0.556
		White	0.653	-0.394	0.897	0.72	0.679	0.669	0.769	0.788	0.948	0.856	0.908	0.498	0.491
	Metric	Latine	0.828	-0.391	0.939	0.833	0.773	0.782	0.736	0.749	0.967	0.617	0.902	0.556	0.718
		Black	0.655	-0.615	0.865	0.785	0.724	0.719	0.724	0.786	0.894	0.798	0.917	0.626	0.583
		White	0.68	-0.502	0.862	0.715	0.724	0.723	0.677	0.825	0.936	0.852	0.922	0.507	0.499
	Scalar	Latine	0.812	-0.372	0.94	0.838	0.772	0.762	0.744	0.73	0.96	0.616	0.904	0.561	0.727
		Black	0.631	-0.604	0.869	0.79	0.716	0.687	0.73	0.776	0.889	0.8	0.919	0.625	0.592
		White	0.667	-0.493	0.867	0.722	0.726	0.711	0.696	0.818	0.936	0.856	0.921	0.507	0.501
Partial Scalar	Latine	0.804	-0.374	0.942	0.840	0.772	0.762	0.745	0.727	0.957	0.626	0.903	0.563	0.731	
	Black	0.617	-0.605	0.869	0.792	0.715	0.684	0.732	0.759	0.862	0.816	0.916	0.629	0.599	
	White	0.658	-0.496	0.869	0.725	0.725	0.711	0.698	0.817	0.932	0.862	0.920	0.509	0.504	

**Table 3*****Teacher Invariance Global Fit Indices***

	Chi-square ( $\chi^2$ )	CFI	$\Delta$ CFI	RMSEA	$\Delta$ RMSEA	SRMR	$\Delta$ SRMR	McDonal d's NCI	$\Delta$ MNCI
Baseline Latine Teacher	$\chi^2(32) = 108.285, p < .001$	0.877		0.165		0.058		0.651	
Baseline White Teacher	$\chi^2(45) = 736.382, p < .001$	0.911		0.110		0.065		0.114	
Baseline Black Teacher	$\chi^2(32) = 129.604, p < .001$	0.863		0.153		0.066		0.687	
Configural	$\chi^2(96) = 330.12, p < .001$	0.884		0.140		0.064		0.734	
Metric	$\chi^2(110) = 348.25, p < .001$	0.882	-0.002	0.131	0.009	0.110	-0.046	0.730	-0.004
Scalar	$\chi^2(124) = 388.84, p < .001$	0.869	-0.013	0.130	0.001	0.104	0.006	0.704	-0.025
Partial Scalar	$\chi^2(123) = 367.27, p < .001$	0.879	0.100	0.130	0.00	0.101	0.003	0.724	-0.019

## STUDY 2 TABLES

**Table 1.**

*Correlations Among Level 2 Variables*

	1	2	3	4	5	6	7	8	9	10	11	12
1. Concept Development	—											
2. Quality Feedback	0.759	—										
3. Language Modeling	0.609	0.754	—									
4. Positive Climate	0.448	0.448	0.371	—								
5. Negative Climate	0.210	0.278	0.349	0.425	—							
6. Teacher Sensitivity	0.396	0.438	0.515	0.620	-0.406	—						
7. Regard for Student Perspectives	0.338	0.316	0.450	0.489	-0.382	0.745	—					
8. Behavior Management	0.396	0.384	0.410	0.595	-0.461	0.606	0.475	—				
9. Productivity	0.381	0.337	0.402	0.537	-0.391	0.556	0.536	0.574	—			
10. Learning Format	0.363	0.352	0.493	0.461	-0.201	0.633	0.578	0.514	0.538	—		
11. Instructional Support	0.864	0.930	0.895	0.467	-0.317	0.506	0.416	0.443	0.417	0.455	—	
12. Emotional Support	0.446	0.468	0.530	0.806	-0.637	0.889	0.841	0.669	0.638	0.609	0.540	—
13. Classroom Organization	0.455	0.428	0.523	0.633	-0.414	0.717	0.638	0.820	0.849	0.831	0.526	0.764

*Note.* All correlations were significant  $p < .001$

**Table 2.**

*Descriptive Statistics*

	Concept Development	Quality Feedback	Language Modeling	Positive Climate	Negative Climate	Teacher Sensitivity	Regards for Student Perspectives	Behavior Management	Productivity	Learning Format	Instructional Support	Emotional Support	Classroom Organization	Social Skills T1	Social Skills T2	Approaches to Learning T1	Approaches to Learning T2
<b>Valid</b>	408	408	408	408	408	408	408	408	408	408	408	408	408	3250	2900	3250	2905
<b>Missing</b>	91	91	91	91	91	91	91	91	91	91	91	91	91	1193	1543	1193	1538
<b>Mean</b>	2.072	2.268	2.454	5.316	1.241	4.635	4.489	5.041	4.925	4.009	2.265	5.300	4.658	15.055	17.223	1.620	1.871
<b>Std. Deviation</b>	0.678	0.721	0.847	0.669	0.455	0.667	0.678	0.745	0.838	0.856	0.672	0.498	0.678	4.958	4.624	.705	.739
<b>Variance</b>	0.460	0.520	0.718	0.448	0.207	0.445	0.460	0.554	0.701	0.733	0.452	0.248	0.459	24.583	21.383	.497	.545
<b>Skewness</b>	0.628	0.899	0.781	-0.094	3.950	-0.290	-0.488	-0.567	-0.403	-0.144	0.792	-0.892	-0.343	-.228	-.445	.236	-.001
<b>Std. Error of Skewness</b>	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	.043	.045	.043	.045
<b>Kurtosis</b>	0.172	1.148	0.643	0.504	26.799	-0.128	0.393	0.129	0.179	-0.535	0.745	2.192	-0.053	-.488	-.537	-.638	-1.007
<b>Std. Error of Kurtosis</b>	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	.086	.091	.086	.091
<b>Minimum</b>	1.000	1.000	1.000	2.330	1.000	2.670	2.000	2.500	2.000	1.750	1.000	2.500	2.330	0	0	0	0
<b>Maximum</b>	4.250	5.000	5.000	7.000	5.670	6.330	6.250	6.750	7.000	6.000	4.560	6.380	6.170	24	24	3	3

108

		Child Race/Ethnicity		
		White	Black	Latine
Teacher Race/Ethnicity	Latine	30	61	622
	White	515	248	326
	Black	98	647	229

**Table 4***Classification Criteria for Classroom Quality Profiles*

	Information Criteria				Measurement Certainty				Ratio Tests				
					Avg. Posterior Probabilities		Class Size (%)						
# of classes	Log Likelihood	BIC	$\Delta$ BIC	AIC	Entropy	Specific	Global	1	2	3	4	LMR	BLR T
1	-4360.476	8841.176		8760.951	1.000	1.000	1.000	100					
2	-3765.370	7717.090	1124.086	7592.740	0.855	.932 .970	0.951	40	60			0.126	<.001
3	-3537.609	7327.692	389.398	7159.219	0.887	.969 .942 .945	0.952	23	23	54		0.090	<.001
4	-3386.027	7090.652	237.040	6878.054	0.909	.961.946 1.0 .936	0.961	23	53	0.004	22	0.129	<.001

Note. LMR = Lo Mendel Rubin Test. BLRT = Bootstrap Likelihood Ratio Test.

**Table 5*****Quality Profiles Across Sample Statistics***

	Profile 1 n (%)		Profile 2 n (%)		Profile 3 n (%)		Total n
Teacher Latine	14	15.91%	23	26.14%	51	57.95%	88
Teacher White	19	11.73%	47	29.01%	96	59.26%	162
Teacher Black	49	36.57%	16	11.94%	69	51.49%	134
Child Latine	160	15.01%	239	22.42%	667	62.57%	1066
Child White	73	14.12%	150	29.01%	294	56.87%	517
Child Black	283	31.94%	161	18.17%	442	49.89%	886
Child Multiracial	30	23.62%	24	18.90%	73	57.48%	127
Child Spanish	151	15.84%	215	22.56%	587	61.59%	953
Child English	380	24.45%	324	20.85%	850	54.70%	1554
Child Female	278	20.72%	281	20.94%	783	58.35%	1342
Child Male	279	21.27%	291	22.18%	742	56.55%	1312
Child Latine DLL	144	15.64%	206	22.37%	571	62.00%	921
Racial/Ethnic Match	314	21.54%	307	21.06%	837	57.41%	1458
Racial/Ethnic Mismatch	335	20.55%	337	20.67%	958	58.77%	1630
No Language Match	491	22.22%	446	20.18%	1273	57.60%	2210
Language Match	105	16.91%	167	26.89%	349	56.20%	621

*Note.* DLL = dual language learner; Profile 1 = moderate emotional support, low classroom organization, and higher negative climate; Profile 2 = high emotional support, classroom organization, and average instructional support; Profile 3 = high emotional support, classroom organization, and lower instructional support.



**Table 6*****Sample Race/Ethnicity Within Quality Profile***

	% of Sample	Profile 1		Profile 2		Profile 3	
		n	%	n	%	n	%
Teacher Latine	19.50%	14	17.07%	23	26.74%	51	23.61%
Teacher White	37.70%	19	23.17%	47	54.65%	96	44.44%
Teacher Black	28.90%	49	59.76%	16	18.60%	69	31.94%
Teacher Race/Ethnicity Total		82.00		86.00		216.00	
Child Latine	39.60%	160	29.30%	239	41.64%	667	45.19%
Child White	20.60%	73	13.37%	150	26.13%	294	19.92%
Child Black	31.80%	283	51.83%	161	28.05%	442	29.95%
Child Multiracial	5.40%	30	5.49%	24	4.18%	73	4.95%
Child Race/Ethnicity Total		546		574		1476	
Racial/Ethnic Match	40%	314	48.38%	307	47.67%	837	46.63%
No Racial/Ethnic Match	60%	335	51.62%	337	52.33%	958	53.37%
Match Total		649		644		1795	

*Note.* DLL = dual language learner; Profile 1 = moderate emotional support, low classroom organization, and higher negative climate; Profile 2 = high emotional support, classroom organization, and average instructional support; Profile 3 = high emotional support, classroom organization, and lower instructional support

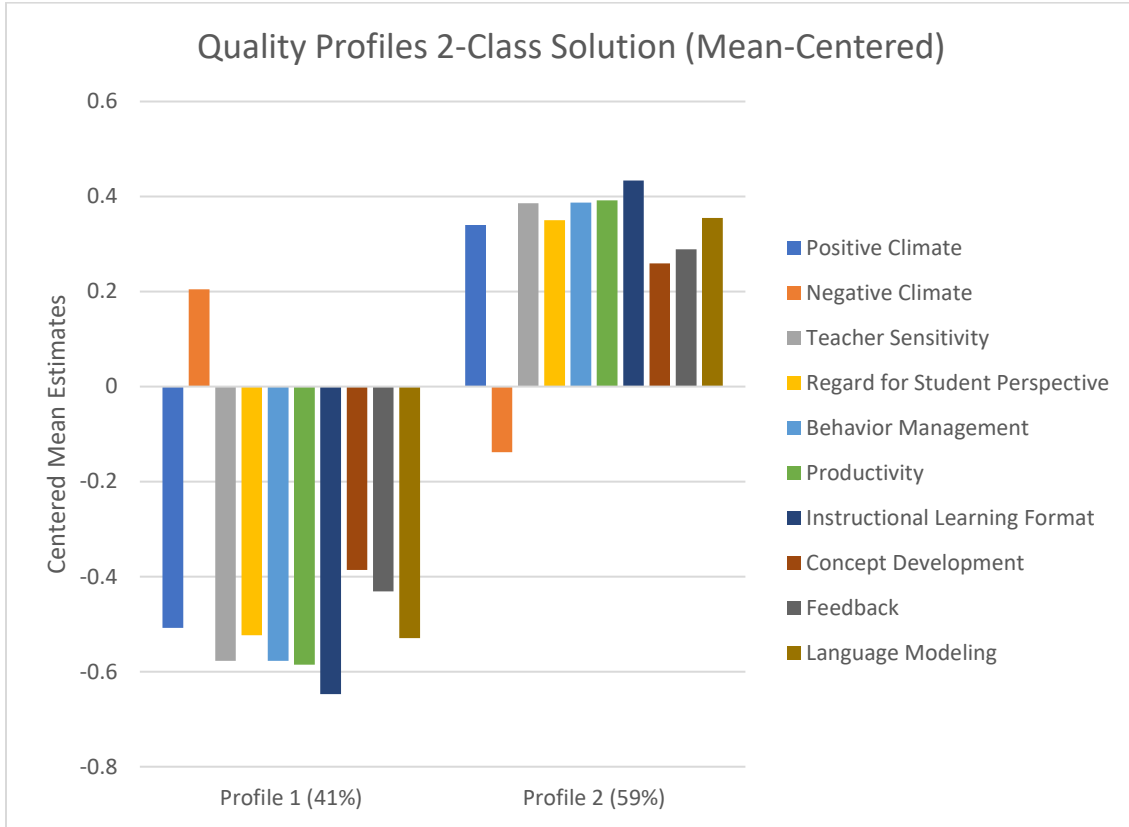
**Table 7*****Quality and Context Predictors of Social Skills and Approaches to Learning***

Predictors	Outcome: Social Skills T2 $\beta$ (SE)			Outcome: Approaches to Learning T2 $\beta$ (SE)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T1	0.547 (0.018)***	0.533 (0.019)***	0.554 (0.018)***	0.624 (0.022)***	0.598 (0.024)***	0.626 (0.022)***
Profile 1	-0.661 (0.314)**	-0.796 (0.324)**	-0.461 (0.305)	-0.124 (0.046)**	-0.164 (0.055)**	-0.106 (0.049)*
Profile 3	-0.304 (0.484)	-0.549 (0.494)	-0.111 (0.469)	0.015 (0.072)	-0.042 (0.077)	0.031 (0.072)
Profile 2	0.116 (0.372)	-0.018 (0.384)	0.305 (0.352)	0.013 (0.051)	-0.021 (0.058)	0.038 (0.052)
Child Black		-0.244 (0.277)			-0.049 (0.049)	
Child Latine		0.430 (0.300)			0.079 (0.049)	
Child Language		-0.038 (0.254)			-0.011 (0.045)	
Child Gender		-0.839 (0.154)***			-0.142 (0.025)***	
Racial/Ethnic Match			-0.211 (0.197)			-0.005 (0.034)
Language Match			0.441 (0.236)			0.045 (0.039)

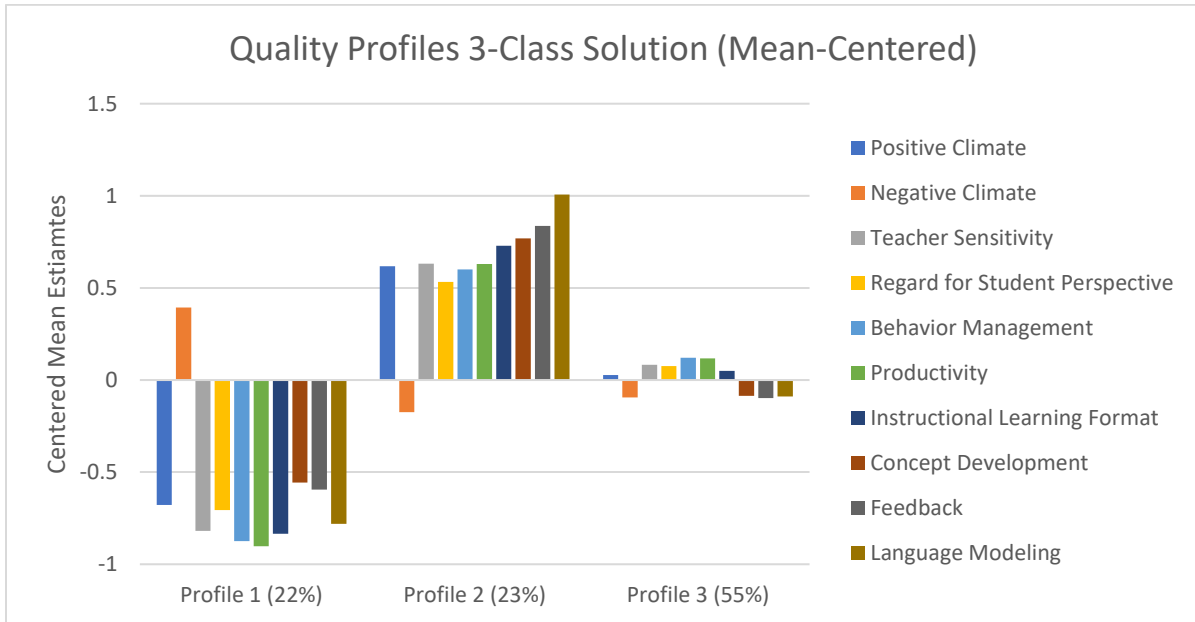
*Note.* \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ . Low-, moderate-, and high-quality refer to posterior probabilities of latent profile membership. T1 = Time 1 (Fall 2009). T2 = Time 2 (Spring 2010). Child Language (0 Spanish, 1 English); Profile 1 = moderate emotional support, low classroom organization, and higher negative climate; Profile 2 = high emotional support, classroom organization, and average instructional support; Profile 3 = high emotional support, classroom organization, and lower instructional support.

APPENDIX B  
STUDY 2 FIGURES

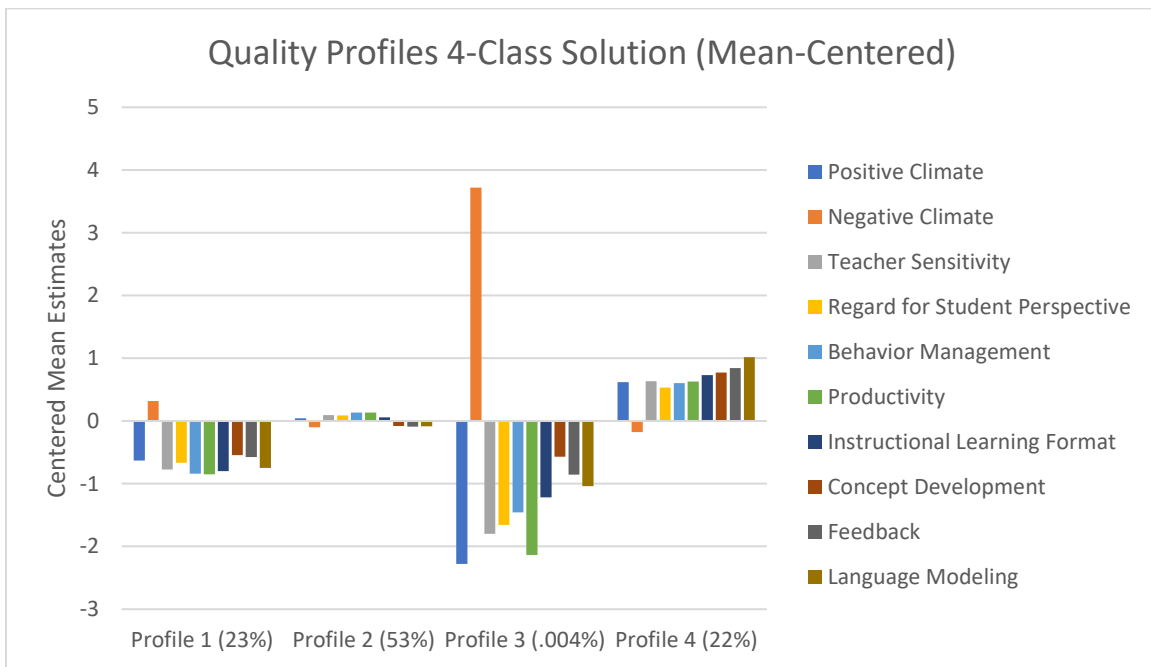
**Figure 1**



**Figure 2**



**Figure 3**



**Figure 4**

