STUDENT-TEACHER RELATIONSHIP QUALITY FOR PRESCHOOL STUDENTS WITH DISABILITIES AND DELAYS: AN EXAMINATION OF STUDENT-TEACHER RELATIONSHIP QUALITY'S MODERATING ROLE AMONGST HEAD START STUDENT, TEACHER, AND CLASSROOM VARIABLES AND KINDERGARTEN READING OUTCOMES

By

Sydney Rae-Saidoo Nelson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

School Psychology – Doctor of Philosophy

2023

ABSTRACT

Student-teacher relationship quality (STRQ) influences a variety of student outcomes, including academic, social, and behavioral functioning (McGrath & Van Bergen, 2015). Less is known about the importance of STRQ for preschool students with disabilities. Many studies are conducted with school-aged, typically developing students (Ansari et al., 2020; Hamre & Pianta, 2001), despite recognition that STRQ may be especially meaningful for young, vulnerable populations (Hajovsky et al., 2017; Sabol & Pianta, 2012). Among studies investigating the importance of STRQ for students with disabilities, most research investigates the connection between student background variables and STRQ, rather than how STRQ may influence other associations (Zendarski et al., 2020). In addition, studies tend to include students with specific disabilities and small sample sizes (e.g., Blacher et al., 2014; Eisenhower et al., 2007), which can be difficult to generalize.

To address these gaps, the Family and Child Experiences Survey (FACES 2009), a nationally representative data set, was used in the present study to examine the following for preschool students with disabilities and delays: 1) the association between student-teacher relationship quality (STRQ) and student's kindergarten academic outcomes, 2) the association between Head Start student, teacher, classroom (HS-STC) variables and student kindergarten academic outcomes while accounting for student background variables, and 3) the role of STRQ as a moderator between HS-STC variables and student kindergarten academic outcomes while accounting for student background variables. STRQ was measured using the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008). Structural equation modeling using latent-factor moderation analysis was completed using *Mplus* Version 8.6 software (Muthén & Muthén, 1998-2017).

Study results suggest that STRQ moderated the association between student gender and kindergarten reading outcomes, such that male students with disabilities and delays performed better on reading outcomes when STRQ in Head Start was strong. Teacher experience and Head Start reading outcomes were significantly associated with kindergarten reading outcomes. Specifically, better kindergarten reading outcomes were associated with students who had Head Start teachers with less experience and students who performed better on preschool reading outcomes. Study limitations and implications for future research and practice are discussed.

ACKNOWLEDGEMENTS

As I complete my dissertation study, I express my sincerest thanks to those who have contributed to my journey to Ph.D. All glory, honor, and praise be given to our precious Lord Jesus. May this dissertation be a blessing to those who read it.

To Dr. Kristin Rispoli, my doctoral advisor and dissertation chair, thank you for your guidance, mentorship, and kindness these past five years. I have appreciated our countless research meetings, your thoughtful manuscript edits, and your continual support as I navigated my doctoral responsibilities and family life transitions. You have helped shape me into an early career psychologist that values science, practice, and scholarship. Thank you.

To my dissertation committee – Dr. John Carlson, Dr. Amy Nuttall, and Dr. Charis Wahman, thank you for sharing your time and expertise, both of which have truly enhanced this dissertation study. To ICPSR, as well as their funders and depositors (i.e., Administration for Children and Families; Mathematica Policy Research; Educational Testing Service; Juarez and Associates), thank you for sharing the FACES 2009 data for this study.

To the FSC Connections lab, thank you for being such an amazing team and contributing ideas during our lab meetings that have propelled this study. To Mackenzie, my cohort-mate and friend, thank you for your support, collaboration, and friendship during this program – I couldn't have done it without you!

To my husband, Bobby, THANK YOU! Thank you for your encouragement during the challenges, your grace during the many hours of writing, and mostly for your unending love and belief in my abilities. You are undoubtedly a reason this dream was possible. To my two little loves, Logan and Avery, I thank you for your joy and energy. Being your mother has shown me the standard of care in which I will strive to treat every child. And, to my "village" who rallied

around me with child-care, a helping hand, and encouragement, this would not have been possible without you.

To my grandparents, thank you for building a strong foundation for our family and modeling the importance of unity, love, and selflessness – all of which I strive to apply to my personal and professional endeavors. Finally, to my parents, David and Janene, I dedicate this dissertation to you. Thank you for your unwavering support and unconditional love. You have taught me the value of hard work, diligence, and perseverance, but most importantly about the gravity of my attitude and character in all situations. I would not be who I am without both of you, and for that – words are not enough to express my gratitude. Thank you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CFA	Confirmatory Factor Analysis
CLASS	Classroom Assessment Scoring System
EST	Ecological Systems Theory
FACES	Family and Child Experiences Survey
HS-STC	Head Start Student, Teacher, Classroom
IDEA	Individuals with Disabilities Education Act
SEM	Structural Equation Modeling
SES	Socio-economic Status
STRQ	Student-Teacher Relationship Quality
STRS	Student Teacher Relationship Scale
WJ	Woodcock-Johnson

CHAPTER 1: Introduction

During the 2019-2020 school year, 7.3 million students ages 3 to 21 years old received special education services as a student with a disability (National Center for Education Statistics, 2021). A disability is a condition that limits a person's ability to interact and engage in day-to-day activities due to impacts on multiple areas of functioning (Center for Disease Control and Prevention, 2020). According to the Individuals with Disabilities Education Act (IDEA, 2004), students can be identified as having a disability in one or more of the following categories: speech or language impairment, intellectual disability, autism, hearing or visual impairment, traumatic brain injury, specific learning disability, orthopedic impairment, other health impairment, or serious emotional disturbance. Students ages three to nine years old may also be considered a student with a disability in school contexts if they experience one or more development delays following an evaluation, such as delays in adaptive, communication, cognitive, physical, social, and/or emotional functioning (IDEA, 2004).

Young students with disabilities and delays experience increased risk for poor academic outcomes. Notable achievement gaps are reported between students with disabilities and their typically developing peers (Stevens et al., 2015). Additional risk factors, such as living in poverty, can exacerbate academic concerns for students with disabilities (Rudasill et al., 2010). Even more, students who live in low-income households are identified for special education at higher rates compared to peers in higher income brackets (National Center for Learning Disabilities, 2020). Students with disabilities are also more likely to demonstrate problem behaviors, be suspended, and have trouble forming strong social connections in school compared to their typically developing peers, which can lead to additional academic challenges (Camacho & Krezmien, 2020; Eisenhower et al., 2007; McGrath & Van Bergen, 2015).

Early education programs and interventions, such as Head Start, aim to support academic success for young children in low-income households, including students with disabilities and delays. Supported by the Individual with Disabilities Act (IDEA, 2004), Head Start requires 10% of students enrolled at each program to be eligible for a disability (Lee & Luellen, 2021; USDHHS, 2018). Head Start programming aims to enhance early learning, child development, and school readiness via instruction, play activities, and high-quality student-teacher relationships. Extant research supports Head Start's emphasis on student-teacher relationships to facilitate student learning as high quality student-teacher relationships are associated with improved academic outcomes (Hamre & Pianta, 2001; Roorda et al., 2011). However, very little is known about the importance of student-teacher relationship quality (STRQ) for preschool students with disabilities and delays.

Theoretical Foundation

Ecological systems theory (EST; Bronfenbrenner, 1979) emphasizes the importance of frequent, often simultaneous interactions that occur across multiple environmental contexts for child development. Each environment, or system, included within EST (i.e., microsystem, mesosystem, exosystem, macrosystem, and chronosystem) underscore how multiple factors in the environment impact developmental trajectories for children (Bronfenbrenner, 1994). The importance of proximal processes, or the ongoing interactions between individuals, on child development guided my examination of STRQ for young students with disabilities as teachers are often an influential adult within students' immediate contexts (Miller-Lewis et al., 2014).

The Process-Person-Context-Time (PPCT) model provides further guidance on the conceptualization of EST (Tudge et al., 2009). The PPCT model includes processes (i.e., proximal processes), personal (i.e., individual characteristics), contextual (i.e., micro-meso-exo-

macrosystems), and time (i.e., chronosystem) as influential factors of development (Tudge et al., 2009). Thus, EST and the PPCT model collectively guided my decision to include Head Start student, teacher, and classroom (HS-STC) variables as predictors of kindergarten reading outcomes as both individual and contextual factors intersect to influence students' development and academic outcomes (Briones et al., 2021; Zatto & Hoglund, 2019) In addition, my study considered how multiple systems simultaneously influenced student academic outcomes by investigating the moderating role of student-teacher relationship quality (STRQ) between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays.

Student-Teacher Relationship Quality (STRQ)

Positive STRQ characterizes student-teacher interactions that are low in conflict and high in closeness, while negative STRQ characterizes student-teacher interactions high in conflict and low in closeness (Rudasill et al., 2010). STRQ influences multiple facets of development, such as student academic, behavioral, and social-emotional outcomes (Decker et al., 2007; McGrath & Van Bergen, 2015; Rudasill et al., 2010). The present study focused on STRQ's association with academic outcomes for students with disabilities. Extant research provides evidence that STRQ influences student academic outcomes for typically developing and at-risk students. For instance, negative student-teacher relationships in early elementary school were predictive of both behavioral and academic outcomes in middle school (Hamre & Pianta, 2001). In a study of behaviorally at-risk youth, improvements in STRQ were associated with improvements in student academic outcomes (Decker et al., 2007). STRQ has also been positively associated with factors related to student learning, such as student engagement (Monahan, 2020).

Researchers of multiple disciplines have increasingly investigated STRQ for typically developing populations in efforts to enhance student outcomes (Sabol & Pianta, 2012). Present

research efforts have heavily reported on the connection between student background variables and STRQ. For instance, students' race tends to be associated with STRQ, as minority students are consistently reported as having more conflictual student-teacher relationships compared to non-minority peers (McGrath & Van Bergen, 2015; Rudasill et al., 2010). Students with lower socioeconomic status (SES) have also been found to experience poorer STRQ than their peers in higher SES brackets (Rudasill et al., 2010). Differences in STRQ based on student characteristics can be a function of racial bias (e.g., Gilliam et al., 2016) and stressors commonly experienced by minoritized students, such as discrimination and system-level barriers to success (e.g., Goldbach & Gibbs, 2017; Mandlawitz, 2003).

STRQ may be especially important for students with disabilities who can experience additional academic, behavioral, and social challenges compared to their typically developing peers (Jimerson et al., 1999; Rudasill et al., 2010). Thus, an important direction for research is to understand how STRQ may serve as a protective factor for students with identified risk factors (Sabol & Pianta, 2012), as improving STRQ could be one way to support students' skill-development (Eisenhower et al., 2007). Unfortunately, existing research suggests that students with disabilities tend to have poorer and less stable STRQ patterns than their typically developing peers (Blacher et al., 2009; Demirkaya & Bakkaloglu, 2015). This is likely due to the unique challenges students with disabilities can experience. For example, compared to their typically developing peers, students with intellectual disabilities can demonstrate increased social and behavioral challenges that result in poorer STRQ (Eisenhower et al., 2007). Beyond evidence that STRQ is poorer for students with disabilities, STRQ research is rather limited, especially for preschool students with disabilities. Learning more about the nuances of STRQ for

students with disabilities can be meaningful when considering practical implications for enhancing student outcomes.

Study Variables

According to ecological systems theory (EST), multiple student and contextual factors influence child development and student outcomes (Bronfenbrenner, 1994). Using EST as a guide, the present study addressed the nuances in how multiple variables together contribute to student academic outcomes by including student, teacher, and classroom components. Specifically, seven Head Start student, teacher, and classroom (HS-STC) study variables were included in the study: a) student gender, b) student race, c) student SES, d) student behavior, e) teacher experience, f) teacher mental health, and g) class size. Reading outcomes were measured by the Woodcock Johnson III Achievement Letter-Word Identification test and served as a proxy measure of academic performance. Finally, student background variables were used as control variables based on research that these variables are associated with student academic outcomes and STRQ. Specifically, two control variables were included: a) student cohort (i.e., completing one vs. two years of Head Start before entering kindergarten) and b) student academic outcomes at entry into Head Start.

STRQ was included as a moderator in my study based on support from EST (O'Connor & McCartney, 2007) and to learn more about how STRQ may strengthen preexisting associations between HS-STC variables and student academic outcomes for students with disabilities and delays. STRQ has also been directly associated with academic outcomes in previous studies (Memon et al., 2019), which was also examined in my dissertation study. The moderating variable (i.e., STRQ) was a latent term using the four dimensions of the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) Emotional Support domain.

The Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) is an observational measure used to assess classroom quality and interactions between teachers and students (Li et al., 2020). Of the three CLASS domains, the Emotional Support domain is closely aligned with a commonly used STRQ informant-report measure, the Student Teacher Relationship Scale (STRS; Pianta & Steinberg, 1992). While rating scales offer unique insight from the teacher's perspective, research suggests that informant reports are subjective as parent, student, and teacher reports do not always align (e.g., Hughes et al., 1999). The CLASS Emotional Support domain was completed by trained, third party raters and was used as a proxy measure of STRQ in this study to reduce potential bias typically associated with informant rating scales.

STRQ as a Moderator

Despite evidence that student, teacher, and classroom variables contribute to student outcomes (e.g., Harding et al., 2019; Mora-Ruano et al., 2019), researchers have neglected to thoroughly consider how STRQ may buffer these associations. This gap in the literature misses potential opportunities to enhance student outcomes as research efforts have been repeatedly dedicated to understanding the connection between student variables and STRQ rather than extending research to understand the role STRQ may have in the association between student, teacher, and classroom variables and student outcomes – especially for young students with disabilities and delays.

Understanding how STRQ can moderate the association between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays lends to targeted intervention opportunities. If STRQ could buffer the association between HS-STC variables and student academic outcomes, focusing intervention efforts on improving STRQ is a streamlined

approach to improving student academic outcomes, rather than dedicating time and financial resources toward multiple intervention efforts aimed at altering HS-STC variables.

Study Aims

Few studies are dedicated to understanding STRQ for preschool students with disabilities and delays as the existing STRQ literature vastly reports on teachers' relationships with school-aged or adolescent students who are typically developing (Eisenhower et al., 2007). My study aimed to add to this notable gap by investigating the following three research questions:

1) Is STRQ formed in preschool significantly associated with kindergarten reading outcomes for students with disabilities and delays?

2) Are HS-STC variables (i.e., student gender, student race, student SES, student behavior, teacher experience, teacher mental health, class size) associated with kindergarten reading outcomes for students with disabilities and delays?

3) Does STRQ in preschool moderate the association between HS-STC variables (i.e., student gender, student race, student SES, student behavior, teacher experience, teacher mental health, class size) and student's kindergarten reading outcomes for students with disabilities and delays?

FACES 2009

The Family and Child Experiences Survey 2009 (i.e., FACES 2009), a nationally representative data set of newly enrolled Head Start students, was used to address these aims. The FACES 2009 data set was chosen as it includes a large nationally representative sample of preschool students with disabilities, which is uncommon in STRQ research. Rather than using national data sets that are inclusive of multiple disability categories, many STRQ studies that include students with disabilities use small sample sizes and include students with specific

disabilities, which can limit the generalizability of findings (e.g., Blacher et al., 2009; Zendarski et al., 2020). In addition, despite the importance of student-teacher relationships early in development (Hamre & Pianta, 2001; Miller-Lewis et al., 2014), many STRQ studies include students in primary or secondary grades, with a dearth of studies including preschool populations (Chen et al., 2018). Collectively, research has failed to merge disability status and early childhood populations when investigating the intricacies of STRQ.

The FACES 2009 study collected data in four waves, between fall 2009 and spring 2012. Data were collected via multiple methods such as direct student assessment, parent and teacher interviews, classroom observations, and surveys. Data pertaining to the student, parent, teacher, home, classroom, and program characteristics were collected. My study used data from all four data collection waves related to student, teacher, and classroom characteristics. Only students identified as having a disability or delay during the first data collection wave were included in the study (n = 367).

Data Analytic Plan

Data analyses were completed using *Mplus* Version 8.6 (Muthén & Muthén, 1998-2017). Data nesting was explored prior to conducting the primary analysis. Frequency, descriptive, and inter-correlational analyses were also completed. Confirmatory factor analyses (CFAs) were conducted for two constructs (i.e., student behavior and STRQ) to examine the factor structure. Structural equation modeling (SEM) was used to test all research questions using one model. To address missing data, the Full Information Maximum Likelihood (FIML) was used. Absolute fit indices were not available for interpretation in the primary model due to the use of a latent interaction in the model (Kenny, 2020). Supplementary analyses were conducted after running the primary model.

Results

A series of confirmatory factor analyses (CFAs) were completed for the student behavior construct prior to conducting the primary analysis. The one-factor CFA with only externalizing behavior items resulted in the best model fit. Thus, a summary score of externalizing behaviors was used as a measure of student behavior in the primary model. In addition, a CFA for STRQ produced adequate model fit. In the primary analysis, STRQ, the variable of interest, was included as a latent term. All other descriptive analyses suggested that normality assumptions were not violated and did not result in multicollinearity.

Research question 1 results indicated that STRQ was not significantly associated with kindergarten reading outcomes for students with disabilities and delays. Research question 2 results suggested that student gender and teacher experience were both significantly associated with kindergarten outcomes. Specifically, identifying as a male student and have a teacher with more experience was associated with poorer kindergarten reading outcomes. All other HS-STC variables (i.e., student race, student SES, student behavior, teacher mental health, and class size) were not significantly associated with kindergarten reading outcomes. One student background variable, Head Start reading outcomes, was significantly associated with kindergarten reading outcomes, such that higher Head Start reading outcomes were associated with higher kindergarten reading outcomes for students with disabilities and delays. Finally, research question 3 results suggested that STRQ significantly moderated the effect between student gender and kindergarten reading outcomes for students with disabilities and delays, such that male students fared better academically when STRQ was high. STRQ did not significantly moderate the effect between any other HS-STC variables and kindergarten reading outcomes for students with disabilities and delays.

Conclusion

My study aimed to understand the intricacies of STRQ in efforts to inform both research and practice related to improving academic outcomes for preschool students with disabilities and delays. The significant moderating effect of STRQ served as a buffer such that male students tended to perform more favorably on kindergarten reading outcomes when STRQ in Head Start was high, which has important implications for future work aimed to improve outcomes for young students with multiple risk factors. Researchers are encouraged to explore alternative statistical models that were beyond the scope of this study, include both a global and individual measure of STRQ, carefully consider the design and study measures, and include diverse samples of students in their studies (e.g., varying disabilities, students enrolled in various preschool programs). Practitioners are encouraged to seek additional training opportunities and implement evidence-based strategies to enhance STRQ in their classrooms, specifically aimed to improve STRQ for male students with disabilities and delays. Finally, researchers and practitioners are encouraged to develop collaborative partnerships to support positive STRQ in schools for students with disabilities and delays.

CHAPTER 2: Literature Review

In this dissertation study, I used the Family and Child Experiences Survey (FACES 2009), a nationally representative data set, to examine the following for preschool students with disabilities and delays: 1) the connection between student-teacher relationship quality (STRQ) and student's kindergarten academic outcomes, 2) the association between Head Start student, teacher, classroom (HS-STC) variables and student kindergarten academic outcomes while accounting for student background variables, and 3) the role of STRQ as a moderator between HS-STC variables and student kindergarten academic outcomes while accounting for student background variables. Substantial evidence points to the importance of STRQ for typically developing students (Hamre & Pianta, 2001; Roodra et al., 2011). My study added to a large literature gap related to understanding the intricacies of STRQ for preschool students with disabilities and delays. STRQ was measured using the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008), an objective, observational measure, as opposed to the commonly used teacher informant-reports in efforts to limit reporting bias.

Theoretical Framework

Ecological Systems Theory (EST; Bronfenbrenner, 1979) emphasizes that child development is impacted by frequent interactions across multiple environmental contexts. Each environment, or system, included within EST (i.e., microsystem, mesosystem, exosystem, macrosystem, and chronosystem) underscore how multiple factors in the environment impact developmental trajectories (Bronfenbrenner, 1994). The microsystem is the closest context in which the individual interacts and engages, such as a student's home or school. The mesosystem encompasses the connection between microsystems that directly pertain to the individual, and the exosystem describes connections between systems, one of which does not directly involve the

individual (yet the interaction influences the individual). The macrosystem encompasses the larger standards of society, such as various beliefs, values, norms, or structures that can affect an individual. Finally, the chronosystem describes the individual or contextual constancy or change that can occur over time (Bronfenbrenner, 1994).

The microsystem primarily guided development of the present study. Bronfenbrenner (1994) emphasized that the organization and opportunities offered within the microsystem are critical to development. Particularly, proximal processes within the microsystem are long-lasting and consistent exchanges between the individual and others, which function to cultivate and bolster development (Bronfenbrenner, 1994). Along with their caregivers, one opportunity for positive proximal processes for students with disabilities and delays occurs with their teacher. Students spend much of their day with teachers, and thus, a teacher's influence increases as preschool students spend more time in the classroom (Hamre & Pianta, 2001; Miller-Lewis et al., 2014). Thus, teachers become an important, consistent individual within the student's immediate context (Miller-Lewis et al., 2014), and can be an influential person within the student's microsystem.

The major construct of interest in the present study, student-teacher relationship quality (STRQ) is conceptualized as proximal processes for students with disabilities that occur within the school context. Young students with positive relationships with teachers tend to experience positive adjustment and social competence (Demirkaya & Bakkaloglu, 2015), as well as improved academic and behavioral outcomes later in development (Hamre & Pianta, 2001). When strong connections with teachers are formed within the microsystem, the student's ability to explore, learn, and grow is enhanced, allowing for comfort to engage in learning. Thus, when students feel close to their teachers, students' attitudes, adjustment, and attendance in school can

all improve, increasing their likelihood of academic success (McGrath & Van Bergen, 2015). Roorda and Kooman (2020) add that positive STRQ helps students feel a sense of emotional security at school, which can increase their likelihood of acclimating positively to the classroom. In turn, a sense of belonging and connection in the classroom provides students with the freedom to take chances and learn new things. Thus, proximal processes between students and teachers in the microsystem can be connected to student outcomes in a variety of ways.

Proximal processes between teachers and preschoolers are especially meaningful as teachers tend to offer emotional support similar to that given by caregivers (Hamre & Pianta, 2001). Even more, preschool students with multiple risk factors (i.e., students with disabilities, low SES) may find the relationship or bond with their teacher particularly important. Students with disabilities and students living in low-income families can experience poorer academic outcomes compared to typically developing peers and students living in higher-income families (Jimerson, 1999; Stevens et al., 2015). Moreover, when students experience multiple risks simultaneously, problems impacting future outcomes can be exacerbated.

Positive proximal processes with teachers can serve as a protective factor amidst other vulnerabilities and school-related challenges. For example, in a study of over 750 elementary school students, students who were academically at-risk only performed well on an academic test when they had positive student-teacher relationships, unlike their typically developing peers who performed well regardless of the relationship with their teacher (Liew et al., 2010). This finding emphasizes EST's principle that a change in one system can impact other systems (Bronfenbrenner, 1979). Using EST, I hypothesized that various HS-STC variables and STRQ could impact academic outcomes for students with disabilities and delays.

The Process-Person-Context-Time (PPCT) model provides further guidance on the conceptualization of EST (Tudge et al., 2009). The PPCT model suggests that processes (i.e., proximal processes), personal (i.e., individual characteristics), contextual (i.e., micro-meso-exo-macrosystems), and time (i.e., chronosystem) all influence development (Tudge et al., 2009). Thus, along with environmental factors, individual characteristics are also purport to impact development (Bronfenbrenner, 1994). Tudge and colleagues (2009) argue to include all aspects of the PPCT model in research, in order to accurately represent EST. My dissertation study embedded aspects of the PPCT and EST. For example, student-teacher relationships (i.e., processes), student factors (i.e., person), classroom factors (i.e., context), and Head Start dosage (i.e., time) were all included in the present study to examine the importance of STRQ for young students with disabilities and delays.

Specifically, the present study included both HS-STC variables and student background variables as guided by EST and the PPCT model. Because EST and the PPCT model emphasize that individual characteristics and multiple systems simultaneously influence student development, this informed my decision to use both student background characteristics and HS-STC variables as exogenous variables. For example, student background variables (e.g., academic performance) and HS-STC contextual variables (e.g., class size) both have implications for student outcomes, particularly for students with disabilities, as higher academic performance and smaller class sizes are both associated with positive academic outcomes (Croninger et al., 2007; Zigmond & Kloo, 2017).

Overall, my study was guided by EST's emphasis on the complexity of child development as it occurs simultaneously amidst multiple individual and contextual factors, which can span across systems (e.g., microsystem, mesosystem, macrosystem, etc.) and time. In

general, the scope of my study addressed the nuances in how various systems simultaneously contribute to student outcomes by examining the moderating role of STRQ between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays. Thus, the effects among multiple variables that cut across varying systems were considered to capture the complexities contributing to student academic outcomes.

Finally, EST also aligns with the Classroom Assessment Scoring System (CLASS) Emotional Support domain, which was used to measure STRQ in the present study. The CLASS measures interactions, or proximal processes, between teachers, students, and peers, as well as the teacher's ability to notice and respond to students' emotional and educational needs (Hamre et al., 2013; Li et al., 2020). Teachers' facilitation of safety, support, and positive interactions in the classroom increases the likelihood of student academic success. As students are provided emotional support from teachers in the classroom, students feel safe to learn, seek help, and explore new ideas (Hamre et al., 2013). Thus, the CLASS Emotional Support dimensions are considered key contributors to forming high-quality STRQ (Rose et al., 2019).

Key Definitions

Students with Disabilities

For my study, teachers reported whether students had a disability based on reports from educational and medical professionals. The medical and educational definition of disability are similar. According to the Center for Disease Control and Prevention (2020), a disability is a condition that limits a person's ability to interact and engage in day-to-day activities due to impacts on learning, memory, physical movement, hearing, vision, communication, thinking, social relationships, and/or mental health. According to the Individuals with Disabilities Education Act (IDEA, 2004), a student with a disability refers to a student who has been

evaluated by a multidisciplinary evaluation team (MET) and needs specialized services to make sufficient educational progress. Like the medical definition, disabilities can include one or more of the following categories: visual or hearing impairment, autism, specific learning disability, intellectual disability, emotional impairment, speech/language impairment, other health impairment, physical impairment, or traumatic brain injury. More broadly, students ages three to nine years old may be considered a student with a disability in school contexts if they experience one or more development delays following an evaluation. Developmental delays can be identified in areas of adaptive, communication, cognitive, physical, social, or emotional functioning (IDEA, 2004). Thus, my study considered a student with a disability or delay based on the medical and educational criteria as teachers received reports from both professionals.

Head Start

Head Start, the United States' biggest federally funded early education and intervention program, prepares infants, toddlers, and preschool students residing in low-income households for school (U.S. Department of Health and Human Services [USDHHS], 2020b). To achieve this aim, Head Start programs prioritize early learning and development, family well-being, and student health. Salient to this study is Head Start's emphasis on early learning and development, particularly through high-quality student-teacher relationships. Head Start services are provided in centers and in homes free of charge for families. Aligned with IDEA requirements that students with disabilities receive a free and appropriate public education, Head Start requires 10% of the total enrollment at each program be students with disabilities (Lee & Luellen, 2021; USDHHS, 2018). Thus, students who receive Head Start services are comprised of diverse racial and ethnic backgrounds, familial situations, and ability-statuses.

There is mixed evidence regarding the short-term impacts of Head Start for students with disabilities, as some studies have found positive outcomes and others have reported little impact of participation in Head Start for students with disabilities (Lee et al., 2015; Lee & Rispoli, 2016). My study further investigated this topic by examining the association between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays using the Head Start Family and Child Experiences Survey (FACES) 2009 study. Notably, it is important to improve clarity about what factors may be especially salient in predicting outcomes for students with disabilities and delays in efforts to advance policy, practice, and research.

FACES 2009

My study used data from the FACES 2009 data set to investigate the importance of STRQ for students with disabilities and delays. The FACES 2009 is a nationally representative data set of three and four-year old students who were newly enrolled in Head Start in 2009. The FACES studies were conducted in 1997, 2000, 2003, 2006, 2009, and 2014 (USDHHS, 2009a). The most recent survey, the FACES 2014, was not selected for this study as the data lack a follow-up data collection period needed to answer the proposed research questions. In the FACES 2009, data were provided related to student and family characteristics, classroom quality and practices, and program characteristics. These data were collected using a variety of modalities, including observations, assessments, surveys, and interviews. Collectively, the FACES 2009 data provide ample opportunity to investigate a variety of research topics related to students, classrooms, and program outcomes (USDHHS, 2009a).

The FACES 2009 is guided by a conceptual framework that emphasizes multiple, interacting relationships can influence developmental outcomes for children (see Appendix A). The student is centrally located in the model and surrounded by the parents and family,

classroom and teachers, Head Start programming, and community, state, and national factors. Like EST, the surrounding factors influence the child's growth, development, and readiness for school. The FACES 2009 conceptual framework aligned with my study, particularly how student-teacher relationship quality (STRQ) intersects with various student, teacher, and classroom variables to influence future academic outcomes for students with disabilities and delays.

Student-Teacher Relationship Quality (STRQ)

STRQ Definition

Student-teacher relationship quality (STRQ) is often described in the literature as either positive or negative. Specifically, high-quality, or positive student-teacher relationships are associated with closeness between the dyad, while low-quality or negative relationships are associated with conflict between the teacher and student (Hughes et al., 2005; Rudasill et al., 2010). Closeness encompasses warmth, care, and respect between the teacher and student. Conflict refers to irritation, frustration, and dissonance among the dyad (Rudasill et al., 2010). However, depending on the informant, closeness and conflict may be perceived differently based on the individual's perception of the relationship dynamic (McGrath & Van Bergen, 2015).

My study took a different, broader approach to measuring STRQ. In the present study, STRQ was conceptualized at the classroom level, rather than focusing on each student's individual relationship with the teacher. Specifically, STRQ was measured by a third-party rater's observation of the emotional support provided to the students in the classroom, namely through the cultivation of a positive classroom environment and the use of sensitive teaching practices (Li et al., 2020). An observational approach offered an objective rating of STRQ via a positive classroom environment and student-teacher interactions.

Although STRQ was measured at the classroom level, it maintained the main definitional components presented in the literature when informant-report measures are used. A positive classroom environment includes strong student-teacher connections or bonds, while a negative classroom environment includes student-teacher conflict or tension, such as hostility, anger, or irritation. Thus, the concept of a positive and negative classroom environment used in this study aligned with the literature's interpretation of closeness and conflict, respectively. Teacher practices, such as their sensitivity and consideration of student views, were also included in my study's STRQ definition. Namely, positive STRQ is cultivated by a teacher who is aware of and responsive to student needs, as well as who embeds student interests and perspectives within classroom practices and procedures.

STRQ Importance

STRQ is particularly salient in students' early education as it has been connected to later academic success, engagement, and adjustment among students (McGrath & Van Bergen, 2015). In a meta-analysis of 99 studies, positive STRQ was positively associated with learning and negative STRQ was negatively associated with learning (Roorda et al., 2011). In younger grades, negative STRQ had a stronger association with academic performance compared to secondary grades. Similarly, in a longitudinal study of 179 students from kindergarten to eighth grade, STRQ high in conflict was negatively associated with academic outcomes in lower elementary, upper elementary, and middle school, with the strongest correlations in lower elementary (Hamre & Pianta, 2001). Thus, STRQ tends to be especially important in early grades, which reinforced my decision to understand the nuances of STRQ in preschool populations in the present study.

STRQ is an essential component of early educational programming that aids students' success in later school years (Pianta et al., 2008), so much so that researchers argue "positive

relationships with adults are perhaps the single most important ingredient in promoting positive student development" (Pianta et al., 2012, pg., 370). Current nation-wide educational initiatives, such as Positive Behavioral Interventions and Supports (PBIS) and the Pyramid Model, integrate STRQ in their approach. PBIS aims to teach and encourage positive behavior while identifying students who need additional support using data-based decision making within a three-tiered system (Center on PBIS, 2022). A foundational practice of PBIS is creating a positive school-wide culture through the cultivation of positive student-teacher relationships. The Pyramid Model is a framework that guides the delivery of evidence-based social emotional supports and interventions in a tiered, scaffolded way, based on student need (NCPMI, 2022). The Pyramid Model prioritizes universal social-emotional supports, such as "nurturing and responsive relationships" that are provided to all students. Overall, both the Pyramid Model and PBIS align with EST as they emphasize the importance of positive relationships for students' future success, as well as consider how to support student development across multiple systems of care. Head Start also emphasizes the development of positive student-teacher relationships by embedding this into their mission. Thus, schools have recognized the importance of STRQ for future academic and behavioral success.

The focus of my study was on student academic outcomes, but it is important to note that STRQ has been associated with multiple school-related outcomes, such as behavioral functioning, attendance, student attitudes, peer relationships, school adjustment, and engagement (McGrath & Van Bergen, 2015). For instance, in a study of 44 elementary students and their teachers, better behavioral and social emotional outcomes were reported by teachers when STRQ was strong (Decker et al., 2007). Hamre and Pianta (2001) report increased rates of suspensions later in elementary school for students with higher student-teacher conflict in kindergarten, even

when student behavior was controlled. On the other hand, close student-teacher relationships were associated with lower problem behaviors later in the school year, particularly for students who entered kindergarten with externalizing behavioral concerns (Silver et al., 2005). In another study of 44 preschool students, student-teacher relationships that were low in conflict were associated with better play skills with peers (Griggs et al., 2009). In addition, low student-teacher conflict was found to be a moderating factor between student temperament and disruptive play with peers ($p < .01$). Thus, multiple studies point to the association between STRQ and behavioral and social outcomes, such that student-teacher relationships that are high in closeness tend to result in positive outcomes for students, while student-teacher relationships that are high in conflict tend to result in the opposite (Blacher et al., 2014).

Though STRQ is connected to multiple student outcomes, my study focused on student reading outcomes. There is well-established evidence pointing to the connection between STRQ and academic outcomes (Ansari et al., 2020). Among a sample of 490 students, Pianta and Stuhlman (2004) report overall student-teacher relationship quality ($\beta = .06$, $p < .01$), closeness ($\beta = .22$, $p < .01$), and conflict ($\beta = -.10$, $p < .05$) predicted teacher reports of student academic performance in first grade. Publications using large data sets echo these findings, such as the study by O'Connor and McCartney (2007), which found that positive STRQ was associated with better student academic performance when using the National Institute of Child Health and Human Development Study of Early Care and Education sample of 1,364 students. McGrath and Van Bergen's (2015) literature review of 12 review articles and 92 studies also presents multiple connections between STRQ and student academic outcomes. For example, STRQ was found to be associated with reading performance in the Program for International Student Assessment 2000 data set comprised of 3,748 high school students across nearly 150 schools (Lee, 2012).

Specifically, when students reported a one unit increase in STRQ, their reading performance increased by nearly 11 units ($p < .001$). Finally, STRQ classroom components (i.e., positive emotional climate, strong teacher sensitivity, emphasis on problem solving, and employing interactive instructional learning opportunities) have also been found to be associated with better academic outcomes for students (Allen et al., 2013).

STRQ is not only important due to its direct association with academic outcomes, but STRQ can also influence other school-related behaviors such as students' attitudes, motivation, and engagement, that are associated with academic outcomes. For example, students tend to have lower motivation to complete academic tasks, often resulting in poorer academic performance, when they report negative STRQ (Hughes et al., 1999; McGrath & Van Bergen, 2015). When students experience negative STRQ, they are also at an increased chance for school avoidance (Murray et al., 2008) and school absences (Davis & Lease, 2007). Repeated time away from classroom instruction and educational content can negatively affect academic outcomes. On the other hand, high STRQ is related to students' positive attitudes toward school (Huan et al., 2012), which is important as negative attitudes about school in kindergarten have been associated with lower academic outcomes in fifth grade (Hauser-Cram et al., 2007). Overall, these findings suggest that STRQ can dampen the effects of variables that influence student academic outcomes, such as students' attitudes and motivation.

Like typically developing students, STRQ and academic outcomes tend to be connected for students vulnerable to risk factors, such as students living in low-income households. For example, in a study of 103 low SES kindergarteners and their teachers, students' academic outcomes were positively correlated with STRQ (Mantzicopoulos, 2005). Similarly, at-risk students in classrooms with high STRQ were found to have better achievement outcomes

compared to their at-risk peers in classrooms with low STRQ (Hamre & Pianta, 2005). In a recent study, Loomis and colleagues (2022) reported that positive STRQ mediated the negative effects of students' inhibitory control on Head Start expulsion for low-income preschool students, which is meaningful as expulsion has adverse effects on academic outcomes. Collectively, these findings are important for the present study as my sample included Head Start students who resided in rather low-income households.

Researchers argue strong STRQ may be especially meaningful for students at-risk for lower academic performance (Fowler et al., 2008; McGrath & Van Bergen, 2015). In a study of 106 students, STRQ was reported to be particularly beneficial for students who had lower academic performance during preschool (Cadima et al., 2010). Specifically, students with lower academic skills at baseline performed better on academic outcomes in elementary school when STRQ was high, unlike their same-age peers with higher academic skills at baseline. STRQ moderated this effect only for students with preexisting academic risks. Even more, STRQ can influence other school-related behaviors that are associated with academic outcomes for students exhibiting risk factors. For example, Silva and colleagues (2011) report that students with low SES and academic risks may increasingly form poorer attitudes about school. Yet, strong STRQ can improve students' attitudes toward school for students with low SES backgrounds (Silva et al., 2011), which can increase the likelihood of academic success.

Collectively, this research points to the importance of STRQ for students with multiple risk factors due to the direct effects between STRQ and academic outcomes, as well as the moderating effects of STRQ on academic outcomes for students with a variety of risk factors. My dissertation study used this evidence to further understand the nuances of STRQ for

vulnerable populations with multiple risk factors, particularly preschool students with disabilities and delays living in economically disadvantaged households.

STRQ for Students with Disabilities

Students with disabilities often experience poorer STRQ than their typically developing peers, as well as less stable STRQ over time (Blacher et al., 2009; McGrath & Bergen, 2015). For instance, Prino and colleagues (2016) compared STRQ among children with Autism Spectrum Disorder (ASD), Learning Disabilities (LD), Attention Deficit Hyperactivity Disorder (ADHD), Down Syndrome, and typically developing peers. Besides students with Down Syndrome, all other students with disabilities had poorer STRQ with teachers than their typically developing peers. Specifically, students with ASD had higher conflict ($F(2,158) = 12.73$; $p < .001$) and lower closeness ($F(2,158) = 17.67$; $p < .001$) with teachers. Students with ADHD also had higher conflict ($t = 7.51$, $p < .001$) and lower closeness ($t = -3.35$, $p < .001$) with teachers compared to typically developing peers. Similar, yet non-significant results were found for students with LD. Analogous findings were reported in a recent study of 360 Portuguese students and their teachers, indicating that students with disabilities had higher conflict and less close relationships with teachers compared to their typically developing peers (Freire et al., 2020). In a study of nearly 100 early elementary school students, STRQ was more stable overtime for typically developing youth as compared to students with ID (Blacher et al., 2009). Student-teacher relationships were also stronger for typically developing youth at ages six (TD = 119.8; ID = 113.0), seven (TD = 118.9, ID = 109.3), and eight (TD = 117.2; ID = 110.4) when comparing mean scores. Finally, Demirkaya and Bakkaloglu (2015) investigated student-teacher relationships among 54 preschool students with disabilities and 54 students without disabilities. Preschool students with disabilities were found to have increased conflict ($p < .01$) and decreased

closeness ($p < .01$) with teachers as compared to their typically developing peers. Collectively, these studies point to a similar pattern in that students with disabilities tend to have poorer STRQ compared to their typically developing peers.

STRQ for students with disabilities may be especially important due to the potential for increased school-related risk factors experienced by these students, such as additional academic needs, poorer social and familial connectedness, and increased internalizing and externalizing behaviors (Eisenhower et al., 2007; Gilmour et al., 2019; McGrath & Van Bergen, 2015). Like typically developing peers, STRQ is associated with student reading scores for students with disabilities (Zendarski et al., 2020). Yet, less is known about the role of STRQ for preschool students with disabilities and delays despite evidence pointing to the importance of student-teacher relationships during early school years (Hughes et al., 2005). Even more, there is a paucity of work that investigates how STRQ may enhance or buffer preexisting associations between various student, teacher, and classroom variables (e.g., student gender, teacher experiences, class size) and student academic outcomes for preschool students with disabilities and delays.

Although not given much attention in the literature, there is some evidence that STRQ may dampen the association between HS-STC variables and kindergarten academic outcomes for students with disabilities. For instance, negative attitudes about school in kindergarten have been associated with lower academic outcomes in fifth grade for students with disabilities living in low-income settings (Hauser-Cram et al., 2007). Positive STRQ can improve students' attitudes toward school (Silva et al., 2011), which may in turn positively influence students' academics. My dissertation study added to this notable gap in the literature by further examining the

intricacies of STRQ for students with disabilities and delays in an effort to understand how STRQ may influence academic outcomes for young students with multiple risk factors.

Measurement of STRQ in the Current Literature

Student Teacher Relationship Scale (STRS). STRQ is commonly measured by a teacher-report measure, the Student Teacher Relationship Scale (STRS; Pianta & Steinberg, 1992). Other STRQ informant-report measures exist, such as the Quality of Student Teacher Relationship Scale (Davis, 2001) and the Teacher-Student Relationship Inventory (TSRI; Ang, 2005), but researchers often use the STRS when examining the STRQ literature (e.g., Blacher et al., 2014; Caplan et al., 2016). The STRS is likely utilized in American studies due to its relatively more well-researched psychometric features (Pianta, 2001). The psychometric properties of other STRQ measures, like the Teacher-Student Relationship Inventory (TSRI), are currently being investigated in other countries (Ang et al., 2020). Item content also slightly differs across the measures as the TSRI includes items related to Satisfaction, Instructional Support, and Conflict, and the STRS items focus on Closeness, Conflict, and Dependency, which may be due to cultural differences in STRQ conceptualization.

The STRS is a 30-item questionnaire and provides scores for Closeness (11 items), Conflict (12 items), and Dependency (4 items). Most STRQ-related studies report the closeness and conflict scores to represent positive and negative aspects of student-teacher relationships, respectively (e.g., Blacher et al., 2014). Raters respond to items using a five-point Likert scale ranging from one (definitely does not apply) to five (definitely does apply). According to the STRS Professional Manual, the STRS has the highest internal reliability scores for the Closeness and Conflict domains [(i.e., Closeness = .86, Conflict = .92, Dependency = .64); (Pianta, 2001)].

Although the STRS is a high-quality, commonly used measure, the consistency of ratings between different stakeholders, such as parents and teachers, is problematic. For example, in a study of 36 students with ASD and their teachers, both parent and teacher STRS data were collected, along with measures of social skills and problem behavior (Blacher et al., 2014). Compared to parent reports, teacher reports resulted in stronger, significant correlations between problem behavior and student-teacher relationships using the STRS, suggesting that teachers and parents may differ in their responses when using informant-reports. Teachers and elementary school children (i.e., second and third grade students) were also found to have differences in reporting of STRQ using the STRS. For example, when teachers and students at-risk for aggression were asked to rate STRQ, the correlations across three years were not significant (Year 1, $r = .23$, Year 2, $r = .15$, Year 3, $r = .11$) (Hughes et al., 1999). Thus, when using informant-report measures, discrepancies between teachers' and students' ratings may emerge (e.g., Zee et al., 2020), which can make it challenging to obtain an objective understanding of STRQ. While these differences may be attributed to difficulty in self-reporting for young students, this finding highlights a flaw in informant-report measures as perceptions may differ across raters.

Along with differences in STRS ratings between parents, teachers, and students, other methodological issues have emerged in the literature. First, there is a concern about the shared method variance and use of a single method approach. For example, when teachers report on relationship quality and other variables, such as the student's behavior, this may lead to reporting bias as all information is being collected from the same rater. An additional concern regarding the STRS is that it does not measure observed behavioral patterns related to STRQ. Rather, the STRS measures the rater's perception of student-teacher behaviors. Thus, an alternative measure

of STRQ was used in my dissertation study to avoid these concerns and to offer another methodological approach to understanding STRQ.

Classroom Assessment Scoring System (CLASS). Authored by the same researchers as the STRS, the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) is an observational measure used to assess classroom quality and interactions between teachers and students (Li et al., 2020). There are multiple versions of the CLASS, including the Infant Version, Toddler Version, Pre-K Version, K-3 Version, Upper Elementary Version, Secondary Version, and the inCLASS [(i.e., individualized CLASS); (University of Virginia, n.d.; Teachstone, n.d.)]. Because there are several versions of the measure, the CLASS can be used in a variety of school-based settings, from early childhood to secondary classrooms, and across multiple subject areas. Of importance to this study, the CLASS has been used in Head Start classrooms across the United States to measure student-teacher interaction quality (USDHHS, 2020a). The FACES 2009 study, which was used in this dissertation, used the CLASS Pre-K Version.

The CLASS includes three domains: Classroom Organization, Instructional Support, and Emotional Support. Within the Classroom Organization domain, there are three dimensions: Behavioral Management, Productivity, and Instructional Learning Formats. The Instructional Support domain also includes three dimensions: Concept Development, Quality of Feedback, and Language Modeling. Finally, the Emotional Support domain includes four dimensions: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. A meta-analysis of the CLASS factor structure confirmed the use of three domains, as proposed by the initial developers (Li et al., 2020).

The Emotional Support domain was selected as a measure of STRQ for this study due to its focus on classroom climate and student-teacher interactions. Emotional Support was used as a proxy for STRQ as researchers suggest the four Emotional Support dimensions contribute to STRQ (Rose et al., 2019). Regarding the Emotional Support dimensions, Positive Climate refers to the teacher-child connection and classroom emotional tone. Negative Climate refers to any negativity within the classroom, such as aggressiveness or irritability. Teacher Sensitivity refers to the teacher's ability to notice and respond to students' academic and emotional needs. Finally, Regard for Student Perspectives refers to whether student-teacher interactions consider students' perspectives and interests (Li et al., 2020). Following a classroom observation, each of these dimensions are assessed by the rater using a one to seven scale, with scores of one and two in the low range, three through five in the middle range, and six and seven in the high range. Higher scores indicate stronger Emotional Support or STRQ.

The Classroom Organization and Instructional Support domain were not selected as a measure of STRQ because these domains do not align with the definition of STRQ. The Classroom Organization domain measures activities associated with the classroom's structure and management, such as the layout of learning centers, stability of schedules and routines, and the overall cohesiveness between staff members (U.S. Department of Health and Human Services, 2021). The Instructional Support domain measures activities associated with the students' language and cognitive development. Specifically, this domain focuses on the strategies teachers use to teach the curriculum, such as relating the educational content to students' lives, scaffolding support, and providing feedback. Thus, the Classroom Organization domain and the Instructional Support domain are broadly focused on classroom management and implementation of academic strategies to support student learning. Though these domains

measure important classroom components, they do not directly relate to understanding STRQ, like the CLASS Emotional Support domain (Moen et al., 2019).

Evidence suggests that the CLASS and STRS are associated, particularly the CLASS's Emotional Support domain and the STRS's ratings of closeness and conflict. In a recent study of 240 students located in high-risk, low-income schools, the observed CLASS Emotional Support domain was positively associated with the teacher-rated STRS Closeness scale ($\beta = .212$, $p = .038$) and negatively associated with the teacher-rated STRS Conflict scale ($\beta = -.268$, $p = .001$; Walker & Graham, 2021). Similarly, the CLASS Emotional Support predicted STRQ using the STRS in a study of 267 preschoolers with developmental delays from low SES backgrounds and their teachers (Moen et al., 2019). Specifically, observed emotional support predicted growth in overall teacher-rated STRQ ($\beta = .59$, $p = .03$) and closeness ($\beta = .97$, $p = .03$) from the fall to late spring of students' first year in preschool, and decreases in conflict ($\beta = -.18$, $p = .03$). These findings indicate that students in classes with better observed emotional support had less conflict with teachers and stronger gains in overall teacher-rated STRQ and student-teacher closeness across the school year. Moen and colleagues' (2019) findings were particularly important for the present study, as my sample included early childhood students with multiple risk factors, specifically an identified disability and delay and low SES.

Overall, the CLASS Emotional Support Domain has been associated with common measures of STRQ, specifically higher observed emotional support has been linked to higher levels of teacher reported closeness. Thus, observed emotional support is connected to STRQ in that higher emotional support aligns with student-teacher closeness and lower emotional support aligns with student-teacher conflict. Collectively, these studies provided support for my use of the CLASS Emotional Support as a proxy measure of STRQ.

Exogenous Study Variables

My study examined the moderating role of STRQ among Head Start student, teacher, and classroom (HS-STC) variables and kindergarten reading outcomes for preschool students with disabilities and delays, while controlling for student background variables. The following HS-STC variables were included as study variables: student gender, student race, student SES, student behavior, teacher experience, teacher mental health, and class size. Based on research with typically developing and at-risk populations suggesting that these HS-STC variables tend to be associated with academic outcomes (e.g., Croninger et al., 2007; Dominguez & Greenfield, 2009; McLean & Conner, 2015; Wei et al., 2011), these specific HS-STC variables were chosen as study variables to further understand how they are associated with future academic outcomes for preschool students with disabilities and delays, as well as to consider how STRQ may influence these associations. The following sections examined the interrelations between the HS-STC study variables, academic outcomes, and STRQ in support of this approach. Due to notable gaps and the dearth of studies that have investigated these constructs explicitly for preschool populations and students with disabilities, much of the content in subsequent sections includes evidence from studies of typically developing students in primary and secondary grades.

Student Gender

Gender differences in disability identification and eligibility has been commonly reported. For example, the National Center for Education Statistics (2021) reports that about 18% of males received special education services in the 2019-2020 school year compared to 10% of female students. In addition to gender differences in disability identification, the type of disability can also differ between males and females. In the 2019-2020 school year, more female

students were identified as having a specific learning disability while more males were identified as having autism (National Center for Education Statistics, 2021).

Gender differences in academic performance have also been reported for students with disabilities. In a study of 3,421 students with disabilities, female students with disabilities were found to have statistically lower performance on the Woodcock Johnson's Letter Word Identification (WJ-LWI) compared to males with disabilities (Wei et al., 2011). This finding is meaningful for the present study as the WJ-LWI was used as an outcome variable to measure student academic outcomes. Gender differences in writing performance have also been reported for males and females at-risk for dyslexia (Berninger et al., 2008).

Females tend to have relationships higher in closeness and lower in conflict compared to their male peers (Jerome et al., 2009; McGrath & Bergen, 2015; Rudasill et al., 2010), which is meaningful as STRQ is connected to student outcomes. For instance, girls with high student-teacher closeness and boys with low student-teacher conflict tend to have better outcomes than their counterparts (Hamre & Pianta, 2001). Boys with high student-teacher conflict experience significant negative academic implications in lower elementary, upper elementary, and middle school. However, for girls, STRQ and academic outcomes are not as strongly correlated in both lower and upper elementary compared to their male peers. Even more, student-teacher conflict was not found to be significantly correlated with academic outcomes in middle school for girls (Hamre & Pianta, 2001). Collectively, these findings are salient to the present study as they suggest that male students tend to have poorer STRQ compared to their female peers, and STRQ for boys tends to be more strongly related to their academic outcomes throughout their education.

STRQ tends to be strongest for both males and females in earlier grades as strong STRQ seems to decrease as students progress through school for both genders. In a longitudinal investigation of the association between STRQ and academic outcomes during elementary school using the National Institution of Child Health and Human Development study sample (n = 1133), conflict increased for males from first to third grade (Hajovsky et al., 2017). Closeness decreased for both males and females from first to third grade and from third to fifth grade, respectively. Similarly, Hamre and Pianta (2001) report that the negative correlations between STRQ conflict and academic outcomes were strongest in earlier grades as compared to higher grades for both males and females. Thus, trends between STRQ and gender have been established. The strongest STRQ for both males and females tend to occur early on, which further supported the present study's investigation of STRQ implications for early childhood populations.

Student Race

Racial disproportionality has commonly been reported in disability identification, particularly for students who identify as African American. For instance, in a literature review of 26 studies, Cruz and Rodl (2018) identified racial disproportionality in special education, as Black students were overrepresented in special education disability categories while Latino students were underrepresented when compared to their white peers. Even more, teacher perceptions of student disability status have been found to differ based on racial category (Cooc, 2017). Compared to their white peers, teachers increasingly associated disability status with minority students, particularly students who identified as Hispanic, Native American, and African American. Thus, compared to other racial groups, racial minority students are at an increased risk of being identified as having a disability particularly in diagnostic categories that

require clinical judgement for diagnosis, such as emotional behavior difficulties, intellectual impairments, and speech and language impairments (Harry & Klinger, 2014).

Along with disproportional representation in special education, racial differences in academic performance have also been reported for students with disabilities. In a study of 3,421 students with disabilities, Black and Hispanic students with disabilities were found to have statistically lower performance on reading outcomes than White students with disabilities (Wei et al., 2011). This gap in academic performance between students of color and their racial majority peers has been a long-standing concern in education. Scholars attribute many system-level factors to this disproportionality, such as cultural mismatch, test biases, educational processes, low income, behavior management strategies, overemphasis on standardized testing, and overall gap in quality education (Skiba et al., 2008; Wiggan, 2007).

STRQ may play a role in the association between students' racial status and teachers' perception of students' academic skills. In a study of over 600 first-grade students who were at-risk academically, relationship qualities were found to mediate the association between a student's African American racial status and the teacher's perception of the child's ability (Hughes et al., 2005). Specifically, positive STRQ improved teachers' perceptions of African American students' academic ability. This study provided initial evidence that STRQ could buffer or enhance the association between student racial status and academic outcomes for students with disabilities and delays.

Teachers' ratings of STRQ tend to be poorer for Black students compared to their White peers (McGrath & Van Bergen, 2015). In a longitudinal study of nearly 900 children from kindergarten to sixth grade, minority students, specifically Black students, were found to be at risk for higher conflict with teachers in kindergarten and had overall higher STRQ conflict scores

from kindergarten to sixth grade compared to their white peers (Jerome et al., 2009). Other studies have echoed these findings in that teachers tend to report higher levels of support for Hispanic and White students as compared to their African American peers (Hughes et al., 2005). The same is true for preschool populations, as teachers reported higher ratings of conflictual STRQ for Black students (Loomis, 2021).

Gaps in teacher-reported conflict scores amongst racial groups tend to increase throughout elementary school. In a study by Jerome and colleagues (2009), high conflict scores increased in an upward linear fashion for Black students overtime compared to their white peers where conflict scores did not increase as rapidly. Although this trend seemed to narrow approaching middle school, the conflict gap between Black and majority students was larger in sixth grade than in kindergarten. Therefore, teachers' perceptions of STRQ with minority students may be lower than racial majority groups in early grades and worsen overtime.

Racial bias may be a contributing factor in differences among teacher ratings of racially diverse students. Multiple studies have demonstrated that teachers tend to rate the behavior of African American students poorer than their white counterparts. For example, in a study of over 500 early elementary students, teachers rated African American students as having more behavioral problems during kindergarten and first grade compared to their white peers (Sbarra & Pianta, 2001). Similarly, using the Early Childhood Longitudinal Study Kindergarten Cohort data, Bates and Glick (2013) found Black students were rated poorer on externalizing measures by white teachers compared to their Black teachers. Gilliam and colleagues (2016) examined early childhood teachers' implicit bias when rating challenging behaviors. One hundred and thirty-five teachers were asked to identify challenging behaviors of white and black female and male students during 30-second video clips. No challenging behaviors were present in the video

clips. When researchers tracked teachers' eye movement, the results indicate that teachers spent more time watching black male students compared to their peers. These results may also be connected to findings that indicate Black students are at a higher risk of receiving a disability in judgement categories as teacher biases may contribute to these diagnoses (Harry & Klinger, 2014). Using this research as support, my study examined the connection between race and kindergarten outcomes for preschool students with disabilities and delays, as well as how STRQ may moderate this association using an objective measure of STRQ.

Student Social Economic Status (SES)

Poverty is a salient factor related to student disability. According to the National Center for Learning Disabilities, students living in poverty are identified for special education at higher rates than their higher-income peers (2020). In 2019, 6.5% of students living in poverty had a disability compared to 3.8% of students living over the poverty line (U.S. Census Bureau, 2021). Even more, SES can influence racial differences in disability diagnosis. Using the Educational Longitudinal Study (N = 10,840), SES explained the disproportionally in LD identification for both Hispanic and Black students compared to their white peers (Shifrer et al., 2011). Thus, racial differences in disability identification were eliminated for these minority groups when SES was accounted for in the analyses.

Students living in low SES households may experience increased risk factors and poorer educational outcomes compared to their higher SES counterparts. For instance, low SES status is often associated with a higher risk of experiencing trauma, stress, and community violence (Burdick-Will et al., 2011; Nelson & Sheridan, 2011). Low SES is also frequently associated with poorer student academic and behavioral outcomes (Jimerson et al., 1999; Rudasill et al., 2010). In a study of 171 students across six school districts, average oral reading fluency scores

were higher in high-and middle-income districts compared to low-income districts (Hoffman, 2018). Moreover, gaps in academic performance have been noticeable and consistent between students in high and low SES categories across multiple decades when researchers aggregate secondary, longitudinal data sets [(N = 2,737,583 students); (Hanushek et al., 2019)].

The connection between SES and academic outcomes is also salient for students with disabilities. In a study of 3,421 students with disabilities, students in the highest SES bracket performed better on the Woodcock Johnson's Letter Word Identification (WJ-LWI) than students in the medium SES bracket (Wei et al., 2011). Wei and colleagues' (2011) findings that students with disabilities in lower income households perform poorer on the WJ-LWI supported my decision to include SES as a study variable as students with similar risk factors were included in my study. It was predicted that STRQ could act as a protective factor for students with disabilities living in low-income households. Similar to Wei and colleagues' (2011), the present study included the WJ-LWI as a measure of student reading outcomes.

The intersection between disability and SES are likely connected to larger system-level barriers. Students in low SES households often attend schools that are financially limited. Multiple barriers are commonly associated with low-income schools, such as less experienced teachers, increased teacher turnover, and fewer resources (Mandlawitz, 2003). Thus, low SES schools experience multiple system level barriers that negatively influences student outcomes (Patterson et al., 2004). This risk is exacerbated for students with disabilities as additional barriers to receiving quality special education evaluations, intervention services, and resources are present. For instance, students with disabilities in low-income districts commonly experience delayed disability identification, delayed access to intervention services, and less parent

involvement (MacArdy, 2008). Without these services and supports, academic outcomes for students with disabilities and delays can suffer.

STRQ may be a way to enhance outcomes for students experiencing low SES as it could break a cyclical pattern of poor SES predicting poorer academic outcomes, leading to lower occupational status and SES post-graduation (McGrath & Decker, 2015). Yet, students in lower SES brackets tend to experience poorer STRQ than students in higher SES brackets. Specifically, students from low-income families can experience more conflict and less closeness with their teachers (Rudasill et al., 2010). In a study of 894 students, family income was negatively correlated with conflict between teachers and students when examining gender, income, and parental involvement on teacher-child conflict. Family income was positively correlated with teacher-student closeness when examining the same factors (Wyrick & Rudasill, 2009). Similarly, in a study of over 1,100 elementary school students, Hajovsky and colleagues (2017) report students with higher SES tend to have lower student-teacher conflict compared to students in lower SES brackets. Finally, the compounded effect of low SES and racial minority status could exemplify the risk for poor STRQ (McGrath & Van Bergen, 2015).

Initial evidence suggests STRQ may buffer the association between SES and academic outcomes for students. In a study by Hughes and colleagues (2005), student-teacher support was significantly associated with students' WJ reading scores ($p < .01$), while parent education (i.e., highly correlated with SES) was not ($p > .05$). Similarly, in a recent study of students with and without ADHD, STRQ was more closely related to students' scores on word reading compared to students' SES (Zendarski et al., 2020). Additional research in this area is important. If STRQ could buffer the association between student demographic variables and student academic outcomes, this has strong implications for interventions as improving STRQ could serve as a

protective factor and indirectly support academic outcomes for students with one or more risk factors.

Student Behavior

Across childhood and adolescence, student externalizing behaviors have long been associated with poorer academic performance (Hinshaw, 1992; Okano et al., 2020). In a recent systematic review of 26 articles, 10 articles demonstrated a significant association between students' externalizing behavior and their academic performance (Kulkarni et al., 2020). Student behavior was important for this study as students with disabilities and delays may experience increased behavioral challenges while at school (Hauser-Cram & Woodman, 2016). The National Center for Education Statistics (2006) reports that up to 40% of students with disabilities demonstrate behavioral concerns at school. This number can vary based on disability type, as it is estimated that approximately 33% of students with speech impairments exhibit problem behaviors compared to approximately 61% of students with emotional impairments (National Center for Special Education Research, 2006).

Behavioral challenges among students with disabilities can also contribute to multiple school-related risk factors that exacerbate the likelihood of poor academic outcomes. For instance, depending on the severity, behavioral challenges can lead to disciplinary infractions, such as out of school suspension. When reviewing statewide data across several years, Krezmien and colleagues (2006) report students with disabilities tend to be suspended at a disproportional rate compared to their typically developing peers. When students are suspended and not in the classroom with their peers, their access to educational content and instruction is limited, which can negatively affect their academic performance. Thus, for students with disabilities who exhibit externalizing behavioral challenges, the negative implications of a suspension can be

compounding as they may already be performing below their typically developing peers (Stevens et al., 2015). When students are in the classroom, externalizing behavioral challenges can hinder a student's engagement, as well as future academic outcomes (Kremer et al., 2016; Olivier et al., 2020).

Similar to school-aged students, research suggests that preschool students' externalizing behavior tends to negatively impact their academic performance in preschool and in future grades (Dominguez & Greenfield, 2009; Spira & Fischel, 2005). Particularly, externalizing behavior concerns in preschool have been associated with delayed reading skills and overall poorer reading outcomes (Campbell et al., 2000; Dominguez & Greenfield, 2009). In a study of 257 Head Start students, Bulotsky-Shearer and colleagues (2011) found that disruptive behavior during structured learning activities was predictive of lower reading scores at the end of the school year. Thus, the connection between behavioral challenges and academic outcomes can be seen as early as preschool.

Externalizing challenges in preschool may also contribute to the heightened rates of preschool expulsion. In a study of Massachusetts preschoolers, the preschool expulsion rate was thirteen times higher than the national expulsion rate for K-12 students, while the preschool suspension rates were less than one fifth the national average for school-aged youth (Gilliam et al., 2006). Thus, unlike K-12 schools, preschools tend to resort to expulsion in response to problem behavior, which is problematic as students are no longer able to participate in foundational learning and socialization activities aimed to prepare them for K-12 schooling. A preschool's response to externalizing behavior can be particularly problematic for students with disabilities and delays who tend to perform lower on national academic tests and may benefit from additional learning supports (National Center for Learning Disabilities, 2017).

When examining STRQ for students with disabilities, most studies investigate how behavior intersects with STRQ. Often, students with disabilities who exhibit high externalizing behaviors exhibit more conflict with teachers. For example, in a study of 36 students with ASD and their teachers, higher externalizing behaviors were positively associated with STRQ conflict scores and negatively correlated with overall student-teacher relationships (Blacher et al., 2014). In a combined sample of typically developing students and students with ID ages six to eight years old, behavior problems were positively related to student-teacher conflict and negatively correlated with the overall student-teacher relationship and student-teacher closeness (Blacher et al., 2009). Despite this trend, studies have found that when teachers exhibit strong emotional support (i.e., measured by the CLASS) students who demonstrate behavioral challenges tend to be less at-risk for poor STRQ (Buyse et al., 2008).

Although research suggests increased behavior concerns are connected to poorer STRQ, STRQ may still temper the negative association between student behavior and academic outcomes. When students with behavior concerns have strong STRQ, they are more likely to perform better on academic tasks. For instance, in a study of students with academic and behavioral concerns, both teacher and student reports of positive STRQ improved students' behavioral outcomes (Decker et al., 2006). Even more, as students reported higher STRQ, academic outcomes also improved. Thus, Decker and colleagues' study exemplifies the importance of STRQ for students with disabilities who do exhibit behavioral concerns, as well as provides evidence that positive STRQ can dampen the effects of behavioral challenges on student academic outcomes.

Teacher Experience

After a thorough search, there was a paucity of research related to how teachers' years of teaching experience influences academic outcomes for students with disabilities. Yet, teacher experience has consistently been correlated with academic outcomes for typically developing students. For instance, as teacher experience increased, student academic scores in math and reading also improved when analyzing data from over 250,000 students. Student academic scores leveled off after teachers acquired about ten years of experience (Ladd & Sorensen, 2017). In another study of over 9,400 students, standardized English scores were significantly related to teacher's experience (Zheng et al., 2020). Finally, using the Early Childhood Longitudinal Data Set – Kindergarten (ECLS – K), having less than two years of experience was negatively correlated with students' first grade reading scores (Croninger et al., 2007). Thus, most studies suggest students who are in classrooms with more experienced teachers have higher academic outcomes.

Multiple factors may contribute to the association between teacher experience and academic outcomes for typically developing students, such as teachers' increased comfort with educational content and delivery when they have acquired more years of experience. Teachers are still learning about their position in the first years in their position (Ladd & Sorenson, 2017). Novice teachers are less experienced in identifying and responding to students' academic concerns and may have fewer tools to address these concerns (Berliner 2001; Berliner 2004; Ladd & Sorenson, 2017). More teaching experience may lead to more awareness and flexibility in responding to academic needs, which in turn can support students' academic success.

STRQ may dampen the effects between teacher experience and student academic outcomes. For example, students' reading ability was more closely related to STRQ (r = .20, p

<.01) compared to teacher's experience (r = .00, p >.05) in a study with and without students with ADHD. This theme was consistent across other academic skills, as students' vocabulary skills and math skills were also more closely associated with STRQ (r = .16, p <.05; r = .14, p <.01), compared to teacher's experience (r = -.04, p >.05; r = -.01, p >.05), respectively (Zendarski et al., 2020). When using the FACES 2014 data set, Head Start student's learning skills in the fall were significantly connected to STRQ (r = .09, p < .01) compared to the non-significant association between student learning skills and the number of years their teacher taught Head Start [(r = .20, p >.05); (Jeon et al., 2021)]. Both studies provided preliminary evidence that STRQ could temper the association between teacher experience and academic outcomes, especially as the samples included students with disabilities (e.g., Zendarski et al., 2020) and Head Start students (e.g., Jeon et al., 2021).

I hypothesized that including teacher experience in the present study was important due to the potential to inform intervention decisions. For instance, if improving STRQ for students with disabilities attenuated the association between teacher experience and academic outcomes, intervention efforts aimed at improving STRQ could indirectly support academic outcomes for young students with multiple risk factors. Thus, rather than waiting for teachers to accumulate years of experience or attempting to rapidly provide additional training or professional development, intervention efforts aimed to improve student-teacher relationships would provide more immediate support for academic outcomes.

Teacher Mental Health

Teacher mental health is important as heightened mental health difficulties and increased stress can lead to teacher burnout. Understanding the implications of teacher mental health on student outcomes is especially prevalent with the influx of mental health challenges associated

with the COVID-19 pandemic. In a recent study of nearly 450 teachers surveyed during the COVID-19 pandemic, teachers reported an average of 7 stressors on an 18-item survey, which was associated with poorer mental health outcomes and barriers to effectively teaching students (Baker et al., 2021). Teacher mental health is also relevant for my study as approximately one third of Head Start teachers reported depressive symptoms across the 2009-2010 academic year using the FACES 2009 data (Hindman & Bustamante, 2019). Consistent with the work of Hindman and Bustamante (2019) using the FACES 2009 data, teacher's depression was used as a proxy for teacher mental health in my dissertation study.

Despite the importance of this topic, research that directly examines the role of teacher's mental health on academic outcomes for students with disabilities and delays is scant. However, teacher mental health has been associated with student outcomes for typically developing and at-risk populations. For instance, lower reports of teacher depression have been linked to positive student wellbeing (Harding et al., 2019). Teacher depression also influences later academic outcomes for students, particularly for students who are at-risk academically (McLean & Conner, 2015). Using the FACES 2009 study, Hindman and Bustamante (2019) found that when teacher mental health improved over the course of the school year, teachers reported improved problem behaviors and prosocial skills for Head Start students. I hypothesized similar findings for students with disabilities in the present study based on convincing evidence presented with typically developing preschool populations, as well as with at-risk populations (Hindman & Bustamante, 2019; McLean & Conner, 2015).

Findings differ when investigating the interrelationship between STRQ and teacher mental health difficulties depending on the STRQ measure (i.e., teacher-report vs. observational measure). In a study of over 200 teachers and 700 students, teachers' depressive symptoms were

not significantly related to teacher's reports of STRQ (Mashburn et al., 2006). Yet, when using observational measures of STRQ, findings differed. In a study of 238 early childhood teachers, raters observed better emotional climates and student-teacher interactions based on teachers' psychological conditions, using the CLASS and ECERS-R, respectively (Pianta et al., 2005). In another study using observational measures, teacher psychological characteristics (i.e., teacher depression, adult-centered attitudes) were significantly related to the classroom's emotional climate and student-teacher interactions. Differences in findings between teacher-report and observational measures may be due to measurement difference based on subjective vs. objective raters. For example, teachers with mental health concerns may have different perceptions of their STRQ compared to an objective rater. The latter studies that include an observational STRQ measure aligned with my study.

The child-caregiver literature provides additional evidence of the connection between mental health and relationship quality. When the Center for Epidemiologic Studies of Depression Scale (CES-D; Radloff, 1997) was used as a measure of depression in a study of over 1200 non-familial caregivers, depression was associated with caregiver-child interactions (Hamre & Pianta, 2004). Specifically, higher caregiver depression was reflective of poorer caregiver-child interactions, mainly less warmth and more withdrawal-like behavior from the caregiver. Similar findings were predicted for early childhood teachers and their students with disabilities and delays, as teacher mental health likely influences STRQ.

I predicted that STRQ would attenuate the association between teacher mental health and student academic outcomes based on research that suggests if teachers can form a strong relationship with students, this may be more meaningful for academic outcomes compared to their own mental health difficulties. For example, academic outcomes are more closely related to

STRQ compared to factors contributing to teacher's stress and mental health challenges such as difficulty of teaching assignment or teaching at-risk students (Mantzicopoulos, 2005). In a recent study of over 1500 Head Start students using the FACES 2014 data set, students' learning skills, such as their attention, independence, and motivation to complete academic tasks was strongly related to STRQ in the fall and spring of Head Start, unlike teacher's depression that was not significantly associated with students' learning skills in the fall or spring (Jeon et al., 2021).

Overall, the topic of teacher's mental health and the connection to STRQ has been understudied, despite evidence to suggest that students can recognize when teachers are feeling stressed (Glazzard & Rose, 2019). Of particular interest in my study was understanding STRQ's moderating role between teacher mental health and student academic outcomes for students with disabilities. Initial evidence suggested that STRQ could dampen the effects of teacher mental health on student outcomes (e.g., Jeon et al., 2021), but further data was needed to address this gap in the extant literature for young students with multiple risk factors. This is important as STRQ is a more malleable factor than teacher mental health. In addition, school interventions that directly aim to improve teacher mental health are rare, unlike initiatives focused on enhancing quality relationships between teachers and students.

Class Size

Understanding how class size contributes to student learning is important as the variable is malleable (Filges et al., 2018). Multiple studies point to the connection between student academic outcomes and the number of students enrolled in the classroom. For instance, using the Early Childhood Longitudinal Data Set – Kindergarten (ECLS – K), larger class sizes were negatively correlated with students' first grade reading outcomes (Croninger et al., 2007). In another study of 21 teachers and over 350 preschoolers, students in smaller classes (n = 15)

demonstrated better literacy skills than their peers in larger classrooms with 20 students (Francis & Barnett, 2019).

Smaller class sizes are beneficial as they can decrease noise levels in the classroom while increasing student-teacher interactions, teacher's instructional feedback provided to students, and students' opportunities to engage in content-related discussions (Zigmond & Kloo, 2017). Specifically for students with disabilities who benefit from specialized instruction, smaller class sizes allow teachers to individualize instruction based on each student's identified need. Thus, the *Handbook of Special Education* argues smaller class sizes are related to better outcomes for students with disabilities (Zigmond & Kloo, 2017). Even more, parents of students with disabilities report that class sizes are a determining factor in their school selection for their child (Mawene & Bal, 2018).

Students in smaller classrooms also tend to have more interactions with their teachers (Folmer-Annevelink et al., 2010). Increased interactions provide opportunities to improve student-teacher relationships and provide additional academic feedback and support. Namely, statistically significant effect sizes were found when comparing student-teacher interactions related to instruction in a class size of 24 students compared to a class size of 15 students in kindergarten (κ = -.70, $p < .05$) and first grade (κ = -1.04, $p < .05$). Similarly, smaller classes had teachers who engaged in more interactions related to classroom management in first grade (κ = -.43, $p <.05$). Thus, interaction frequency regarding instruction and classroom management between students and their teacher increases in smaller classes.

I predicted that STRQ may attenuate the association between Head Start class size and student's kindergarten reading outcomes. Research suggests that if teachers can form high-quality relationships with students, it may be more important than having less frequent

interactions in a large class size. The parent-child literature suggests that the quality of the relationship is more important than the quantity. For instance, using the Early Childhood Longitudinal Study, Zhang (2020) reports that the quality of parent-child communication was significantly related to student academic outcomes, while the quantity of the parent-child communication was not significantly associated with academic outcomes. Moreover, strong STRQ may be more meaningful for academic outcomes even if students have less frequent interactions with teachers when placed in a large class size.

Endogenous Variable

Academic Performance

In the present study, kindergarten reading was used a proxy for academic performance due to the connection between reading abilities and multiple cognitive processes as well as later academic performance. In a recent meta-analysis of 378 studies, Peng and colleagues (2022) reported that reading difficulties are connected to challenges in variety of cognitive processes, like visuospatial skills, executive functioning, and processing speed. Cognitive abilities for students with reading challenges also tend to worsen as students get older (i.e., lower IQ scores, lower reading scores). Even more, Benischek and colleagues (2020) found that better pre-reading skills were associated with better connectivity in the reading areas of the brain among students ages two to five years old. Like the connection with cognitive processes, reading skills in early childhood tend to be predictive of future academic outcomes for a variety of readers, such students who are high, average, and below-average reading performers (Ozernov-Palchik et al., 2017).

Extant early childhood literature has also used reading skills to represent academic performance. For instance, in a study investigating predictors of student-teacher relationships,

Jerome and colleagues (2009) used the Woodcock Johnson (WJ) Achievement to measure academic ability for children (i.e., 54 months old) at school entry. In another study examining predictors of academic skills from preschool to elementary school, the WJ Achievement Letter-Word Identification subtest was used as a measure of reading achievement at four different time points (Burchinal et al., 2002). Finally, in a study investigating STRQ between teachers and kindergarten students, the WJ Achievement was also used as a proxy for academic achievement (Mantzicopoulos, 2005).

Aligned with the early childhood literature, the Woodcock-Johnson III Tests of Achievement-Third Edition (WJ-III) Letter-word Identification was used as a measure of students' reading ability in my dissertation. The Letter-word Identification subtest is a standardized direct assessment that asks children to identify letters and words from a test book. This subtest was chosen for multiple reasons. First, the WJ-III is direct assessment administered to students by a trained professional, which is advantageous as it provides a direct representation of each student's skill, rather than a teacher-report that is subject to informant bias. Second, the WJ-III is a reliable and valid standardized measure and has been used in numerous early childhood studies investigating academic outcomes for students (e.g., Burchinal et al., 2002; USDHHS, 2009a). The WJ-III also measures foundational reading skills (i.e., identifying letters and words) that preschool and kindergarten students will need to develop literacy skills. Specifically, the WJ-III Letter-word Identification task measures phonics, which teaches students how written text and spoken sounds are linked. According to the National Institute of Health's (2020) National Reading Panel report, phonics is a key component of learning to read, especially for students with disabilities. Finally, the WJ-III was normed using a nationally representative

sample of students, which is inclusive of my study's sample of students with multiple risk factors throughout preschool and kindergarten.

The WJ-III was used as the endogenous variable in the study. Specifically, the WJ-III assessment data from the spring of students' kindergarten year (i.e., spring 2011; spring 2012) was used as an endogenous variable in the model. Along with support from the early childhood literature that uses the WJ-III as a proxy for academic outcomes (e.g., Jerome et al., 2009), the WJ-III was administered in the spring of students' kindergarten year in the FACES 2009 study (USDHHS, 2009a). Data collection in the spring provides a better representation of students' academic performance in kindergarten compared to a measure collected in the fall, as students had additional time to develop reading skills in the kindergarten classroom. The timing of data collection was important to fully understand if preschool variables and STRQ were associated with kindergarten academic performance.

Academic performance was also an exogenous control variable. Specifically, the WJ-III assessment data from students' first year of Head Start (i.e., fall 2009) was used as an autoregressive control. Research suggests that prior and current academic performance is predictive of future performance (Croninger et al., 2007). For example, in a study of over four thousand middle school students, academic performance in middle school was a strong predictor of high school academic performance (Casillas et al., 2012). Similar findings are true for early childhood populations who attend preschool. In a recent study of students enrolled in Head Start programs, academic growth (i.e., vocabulary skills) during Head Start was predictive of better kindergarten academic outcomes (Ramsook et al., 2020). Duncan and colleagues (2007) emphasize this connection when examining six longitudinal data sets. After controlling for social emotional skills, attention, and cognitive functioning, academic performance was a strong

predictor of future academic outcomes. Specifically, reading and math performance predicted future academic outcomes better than students' behavior, attention, and social skills (Duncan et al., 2007).

Exogenous Control Variable

It was important to understand the moderating role of STRQ among the association between Head Start student, teacher, and classroom (HS-STC) variables and kindergarten reading outcomes above and beyond the influence of student background variables. Two student background variables were controlled, specifically student cohort and academic performance at entry into Head Start, as described in the previous section. These variables were considered "student background variables" as each variable was unique to each student's background upon entering Head Start, such as their exposure to preschool (i.e., cohort) and performance-level (i.e., academic performance). The following section examines the interrelations between the student cohort and academic outcomes in support of this approach.

Student Cohort

Student cohort represented whether a student spent one or two years in Head Start prior to entering kindergarten. Student cohort needed to be accounted for in my study as differences in students' kindergarten academic outcomes based on Head Start program dosage have been reported. For instance, in a study by Domitrovich and colleagues (2013), students who attended Head Start for two years performed better academically in kindergarten compared to students who only attended one year. This is echoed in other studies that have found students who attended Head Start for two years performed better socially and academically in kindergarten compared to students enrolled in Head Start for one year (Wen et al., 2012).

Although extant research has prioritized samples of typically developing students, similar trends may be present for students with disabilities and delays. Preschool has long been considered a form of primary prevention that is connected to positive student outcomes. Preschool as a form of primary prevention is important for students with disabilities and delays as Head Start programming exposes children to academic materials, as well as other skill building activities that supports academic success. For example, Head Start students can learn social skills and emotion regulation strategies, receive access to medical services, and build relationships between students, teachers, and caregivers (U.S. Department of Health and Human Services, 2020b). Research suggests that social skills (Caemmerer & Keith, 2015), emotion regulation (Graziano et al., 2007), and school-family-community partnerships (Smith et al., 2020) are all interrelated to academic outcomes. Even more, exposure to in-depth reading and math content during preschool is beneficial for kindergarten learning (Claessens et al., 2014). Thus, increasing time spent in Head Start not only lends to additional opportunities to receive academic instruction, but it also exposes students with disabilities and delays to multiple protective factors that are associated with academic success.

Variable Summary

Academic outcomes and STRQ are associated with multiple student background and HS-STC variables. With most attention given to typically developing students in grades K-12, less is known about HS-STC variables for early childhood populations with identified disabilities and delays. My study aimed to fill this gap by examining the following for students with disabilities and delays: 1) the connection between STRQ and student's academic outcomes, 2) the association between HS-STC variables and student kindergarten academic outcomes while controlling for student background variables, and 3) the role of STRQ as a moderator between

various HS-STC variables and student kindergarten academic outcomes while controlling for student background variables.

Gaps in Existing Research

Gaps in understanding how STRQ may moderate the association between various HS-STC variables and kindergarten academic outcomes among students with disabilities or delays are evident, despite recognition that strong STRQ may be especially meaningful for these populations (Sabol & Pianta, 2012). For students with disabilities, most research investigates the connection between student background variables and STRQ, rather than how STRQ may influence other associations. Additionally, most STRQ research tends to neglect the inclusion of preschool populations with disabilities (e.g., Roorda & Koomen, 2021). Even in studies that broadly identify at-risk populations, disability status is often overlooked (Sabol & Pianta, 2012). Among STRQ studies that do include disability status, students are often enrolled in primary or secondary grades rather than early childhood settings (e.g., Zendarski et al., 2020). These studies tend to have a small sample size and include a specific disability category, which poses barriers to generalizing the findings to broader populations (e.g., Blacher et al., 2014; Eisenhower et al., 2007).

Other notable gaps in the literature relate to STRQ measurement and study methods. Measurement of STRQ tends to be collected via teacher report even though studies have found inconsistent reports of STRQ between ratings provided by teachers, students, and parents (e.g., Hughes et al., 1999). Finally, related to the sample and timing of data collection, little is known about the role of STRQ during preschool on kindergarten outcomes for special populations as most studies are conducted across elementary school and secondary school (e.g., Hamre &

Pianta, 2001). Thus, there are multiple gaps in the STRQ literature related to sample, measures, and method that I attempted to address in my study.

Present Study

The aim of the present study was three-fold. First, research question one examined if there was an association between STRQ established in Head Start and kindergarten academic outcomes for students with disabilities and delays when using an objective, observational measurement tool for STRQ. Second, research question two considered the association between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays while controlling for student background variables. Finally, research question three examined the moderating role of STRQ in preschool between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays while controlling for student background variables.

Research Questions and Hypotheses

Aligned with the FACES 2009 study objectives (USDHHS, 2009a), the present study examined the moderating role of STRQ amongst the association between HS-STC variables and kindergarten academic outcomes for students with disabilities and delays. My study aims were addressed through the following research questions and hypotheses:

Research Question 1

Is STRQ formed in preschool significantly associated with kindergarten reading outcomes for students with disabilities and delays?

Hypothesis 1. STRQ in preschool will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between STRQ and kindergarten reading outcomes.

Hypothesis 1 was supported by the general education literature that indicates a connection between STRQ and academic outcomes for typically developing students (e.g., O'Connor & McCartney, 2007; Pianta & Stuhlman, 2004). For example, in a sample of nearly 3,800 students, STRQ was associated with reading performance (Lee, 2012). This was echoed in studies that have used the CLASS as a measure of STRQ. After controlling for prior academic performance, higher CLASS ratings of positive emotional climate and teacher sensitivity (i.e., STRQ) were associated with better academic performance in a sample of nearly 650 students (Allen et al., 2017). Research question 1 added to the existing body of literature using an observational measure of STRQ, rather than the commonly used rating scales completed by teachers, parents, and/or students that may include informant bias.

Research question 1 also contributed to the few studies investigating the connection between STRQ and academic outcomes for diverse student groups. Most studies point to the connection between STRQ and academic outcomes for typically developing students (Hamre & Pianta, 2001). However, there is evidence that STRQ is associated with academic outcomes for vulnerable student groups, including students with multiple risk factors, such as students who are academically at-risk (Fowler et al., 2008) and living in low-income families (Mantzicopoulos, 2005). Thus, like other student populations with risk factors, it was hypothesized that STRQ would be meaningful for the academic success of students with disabilities and delays.

Finally, research question 1 contributed to even fewer studies examining this topic for preschool students with disabilities and delays. STRQ is especially meaningful during early childhood (Demirkaya & Bakkaloglu, 2015) and for at-risk students (e.g., McGrath & Van Bergen, 2015). For preschool students with disabilities, this topic is especially meaningful as increased academic, social, and behavioral challenges may be present (Eisenhower et al., 2007;

McGrath & Van Bergen, 2015), and STRQ could act as a protective factor for students with a variety of risk factors (Sabol & Pianta, 2012).

Research Question 2

Are HS-STC variables associated with kindergarten reading outcomes for students with disabilities and delays?

Hypothesis 2a. Student gender will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between student gender and student kindergarten reading outcomes. Hypothesis 2a predicted student's gender would be associated with kindergarten reading outcomes for students with disabilities and delays. Male students tend to receive special education services at higher rates compared to their female peers, specifically in disability categories such as autism (National Center for Education Statistics, 2021). Hypothesis 2a was supported by the vast literature that suggests that female students tend to outperform male students on reading achievement tests (e.g., Liederman et al., 2005; Logan & Johnston, 2010). However, some evidence points to female students with disabilities performing lower on academic outcomes compared to their male peers (Wei et al., 2011). Based on the majority of research, a positive statistically significant parameter estimate between student gender and student's kindergarten reading outcomes was predicted.

Hypothesis 2b. Student race will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between student race and student kindergarten reading outcomes. Hypothesis 2b predicted student's racial status would be associated with kindergarten reading outcomes for students with disabilities and delays. Minority students, specifically Black students,

tend to be overrepresented in the disability population (Cruz & Rodl, 2018). Hypothesis 2b was supported by research findings that Black minority students with disabilities tend to perform lower on academic outcomes than their white counterparts (Wei et al., 2011). Differences in academic performance and possible overidentification can be attributed to racial biases and system-wide barriers, especially for students with multiple risk factors (Skiba et al., 2008; Wiggan, 2007). Thus, supported by previous research, a positive statistically significant parameter estimate between student racial status and student's kindergarten reading outcomes was predicted.

Hypothesis 2c. Student SES will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between student SES and student kindergarten reading outcomes. Hypothesis 2c predicted student's SES would be associated with kindergarten reading outcomes for students with disabilities and delays. Similar to students who identify as a racial minority, students living in low-income households are disproportionally identified as having a disability (National Center for Learning Disabilities, 2020). In addition, numerous studies point to the connection between poverty and academic outcomes. Namely, students living in low-income households are at greater risk for poorer educational outcomes, as well as risk factors related to academic failure, such as experiencing trauma or community violence (Burdick-Will et al., 2011; Jimerson et al., 1999; Nelson & Sheridan, 2011).

These findings are consistent in the disability literature. For example, in a study by Wei and colleagues (2011), students with disabilities who lived above the poverty line performed better on the WJ Achievement test than students who were living below the poverty line. System level barriers, such as increased teacher turnover, delayed special education evaluations, and

limited resources likely contribute to the academic gaps based on SES, especially for students with disabilities and delays (MacArdy, 2008; Mandlawitz, 2003). Thus, supported by previous research, a positive statistically significant parameter estimate between student SES and student's kindergarten reading outcomes was predicted.

Hypothesis 2d. Student behavior in preschool will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a negative statistically significant parameter estimate between student behavior ratings in preschool (i.e., lower behavior ratings indicative of positive behavior) and student kindergarten reading outcomes. Hypothesis 2d predicted student behavior in Head Start would be associated with reading outcomes in spring of their kindergarten year. Hypothesis 2d was supported by findings that students with disabilities can experience increased behavioral concerns at school (Hauser-Cram & Woodman, 2016). For instance, the National Center for Special Education Research (2006) reports approximately one third of students with a speech impairment exhibit behavioral concerns at school. Behavioral concerns can increase to over 60% for students with an emotional impairment.

Based on studies connecting behavior problems to poorer academic outcomes for typically developing students (Oliver et al., 2020), I predicted that behavior problems in Head Start would be negatively associated with kindergarten academic outcomes for students with disabilities and delays. Behavior problems can lead to risk factors that negatively affect academic performance. For instance, externalizing and internalizing behaviors can interrupt student's engagement in academic tasks and increase the likelihood for disciplinary infractions (Kremer et al., 2016; Krezmien et al., 2006; Olivier et al., 2020). Students with disabilities are suspended at higher rates than their typically developing peers following disciplinary infractions (Krezmien et

al. 2006). Thus, supported by previous research, a negative statistically significant parameter estimate between preschool student behavior and student's kindergarten reading outcomes was predicted.

Hypothesis 2e. Teacher experience in preschool will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between preschool teacher experience and student kindergarten reading outcomes. Hypothesis 2e predicted teacher experience in Head Start would be associated with students' reading outcomes in spring of their kindergarten year. Although minimal research has been collected for students with disabilities and delays, hypothesis 2e was supported by evidence from samples of typically developing students indicating teacher characteristics and student outcomes are related. For instance, when analyzing data from over 250,000 students, student academic scores in math and reading improved as teacher experience increased (Ladd & Sorensen, 2017). This was echoed in a recent study of over 9400 Chinese students, as English test scores were significantly related to teacher's experience (Zheng et al., 2020). In general, students' academic success was predicted to be attributed to teachers' increased comfort with educational content and delivery when they acquired more years of experience.

I predicted a similar connection between teacher experience and reading outcomes would result for students with disabilities and delays. Specifically, students with disabilities and delays who had more experienced Head Start teachers were predicted to perform better on kindergarten reading outcomes compared to students with less experienced teachers. Teachers with less experience may have more difficulty identifying and responding to student's academic needs as they have had fewer opportunities to do so simply based on fewer years working in the

classroom (Berliner 2001; Berliner 2004; Ladd & Sorenson, 2017). Thus, preschoolers with more experienced Head Start teachers were predicted to have better academic outcomes that school year, as well as future years as the benefits would likely carry over into subsequent years. Students' past and current academic performance has been found to be a strong predicter of their future academic performance (Croninger et al., 2007; Duncan et al., 2007). Thus, supported by previous research, a positive statistically significant parameter estimate between preschool teacher's experience and student's kindergarten reading outcomes was predicted.

Hypothesis 2f. Teacher mental health will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a negative statistically significant parameter estimate between preschool teacher mental health difficulties and student kindergarten reading outcomes. Hypothesis 2f predicted students with Head Start teachers with higher rates of mental health concerns would perform poorer on reading measures in the spring of their kindergarten year. Specifically, it was predicted that a negative statistically significant parameter estimate would result between Head Start teacher's mental health difficulties and student's kindergarten reading outcomes. Hypothesis 2f was supported by evidence in typically developing populations that lower teacher depression scores are associated with better student wellbeing (Harding et al., 2019). Similarly, for students who are academically at-risk, higher teacher depression scores were found to be associated with poorer student academic outcomes (McLean & Conner, 2015). Like hypothesis 2e, it was predicted that teacher characteristics that are associated with preschool academic performance would influence kindergarten academic outcomes (Croninger et al., 2007).

Despite a dearth of studies in the disability and early childhood literature, this topic is important for these groups. For instance, one in three teachers endorsed some mental health

concern in wave one of the FACES 2009 (Hindman & Bustamante, 2019). Common predictors of early childhood teachers' mental health difficulties include work conditions (i.e., chaos in the classroom) and student characteristics [(i.e., students' social skills, behavior); (Hindman & Bustamante, 2019; Jeon et al., 2018)]. Although the connection between student's disability status and teacher's mental health is not fully understood, students with disabilities may exhibit increased behavioral and social challenges (Eisenhower et al., 2007), which may be connected to teacher's mental health. For instance, Head Start teacher reports of student problem behaviors and prosocial skills improved when teacher's mental health concerns decreased, emphasizing the connection between teacher mental health and student outcomes (Hindman & Bustamante, 2019).

Hypothesis 2g. Class size will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a negative statistically significant parameter estimate between preschool class size and student kindergarten reading outcomes. Hypothesis 2g predicted that students in larger Head Start classrooms would have lower kindergarten reading scores. Conversely, students in smaller Head Start classrooms would have higher kindergarten reading outcomes. Hypothesis 2g was supported by findings that classroom environmental characteristics are connected to student outcomes. For instance, the *Handbook of Special Education* suggests smaller class sizes are connected to better outcomes for students with disabilities (Zigmond & Kloo, 2017). Smaller class sizes can allow for more one-on-one instruction and feedback, as well as affect the pace and amount of academic material that is taught to students (Zigmond & Kloo, 2017). Teachers with smaller class sizes may be able to recognize the individual needs of students, which can subsequently inform their teaching practices. Research also indicates students with disabilities who are in classrooms with fewer

students tend to have better reading skills compared to their peers in larger classrooms (Francis & Barnett, 2019). Class size in Head Start is meaningful for future academic success as students' academic performance in earlier grades tend to be connected to their future academic outcomes (Croninger et al., 2007; Ramsook et al., 2020). Thus, I predicted that the academic benefits from a small class size in preschool would carry over in subsequent years as indicated by a statistically significant parameter estimate between Head Start class size and kindergarten reading outcomes for students with disabilities and delays.

Research Question 3

Does STRQ in preschool moderate the association between HS-STC variables and student's kindergarten reading outcomes for students with disabilities and delays?

Hypothesis 3a. STRQ in preschool will moderate the association between student gender and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student gender and kindergarten reading outcomes when strong STRQ is present. Females with disabilities tend to experience poorer academic outcomes compared to their male peers (Wei et al., 2011). However, Hypothesis 3a predicted that STRQ would dampen the association between student gender and kindergarten reading outcomes. Preliminary evidence suggested that STRQ may influence the association between student gender and academic ability. For example, in a study of students with and without disabilities by Zendarski and colleagues (2020), STRQ had a stronger association with students' reading scores compared to students' gender. Hypothesis 3a is salient as STRQ is a more malleable variable compared to student gender. If STRQ could buffer the association between student gender and student academic outcomes, this would have strong

implications for interventions as improving STRQ could indirectly support academic outcomes for students with disabilities and delays.

Hypothesis 3b. STRQ in preschool will moderate the association between student race and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student race and kindergarten reading outcomes when strong STRQ is present. Unfortunately, for Black students with disabilities, racial minority status is associated with poorer academic outcomes (Wei et al., 2011). Hypothesis 3b predicted that STRQ would dampen the association between student racial status and kindergarten reading outcomes. Preliminary evidence suggested that STRQ may influence the association between racial status and teacher reports of academic ability. In a study by Hughes and colleagues (2005), STRQ mediated the correlation between Black minority status and teacher perception of student's academic ability. The study consisted of students who were at-risk academically, providing further evidence that STRQ may be particularly important for students with disabilities and delays. Hypothesis 3b is important as STRQ is a malleable factor that can be improved through evidence-based interventions, while student's racial status is fixed. Thus, if STRQ could dampen the negative association between fixed variables and academic outcomes, this could inform future research, practice, and interventions aimed at improving academic outcomes, particularly for young students with multiple risk factors.

Hypothesis 3c. STRQ in preschool will moderate the association between student SES and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student SES and kindergarten reading outcomes when strong STRQ is present. Students with disabilities located in low SES settings can experience poorer academic outcomes compared to students in high-income settings

(Wei et al., 2011). STRQ was predicted to moderate the association between student SES and kindergarten reading outcomes. Specifically, it was hypothesized that positive STRQ would dampen the effects of SES on students' kindergarten reading outcomes when student background variables were controlled as observed by a statistically significant interaction effect between student SES and STRQ.

Hypothesis 3c was supported by extant literature reporting the importance of STRQ for student academic outcomes (e.g., Decker et al., 2007; Hamre & Pianta, 2001; Pianta & Stuhlman, 2004). Preliminary evidence also indicated that STRQ may be more important for academic outcomes compared to SES. For example, in a study by Zendarski and colleagues (2020), STRQ had a stronger association with students' reading scores compared to students' SES. These findings are particularly meaningful for my study as Zendarski and colleagues' (2020) sample included students with and without disabilities, which provides evidence that STRQ is important for academic outcomes for at-risk students. Overall, hypothesis 3c is salient as STRQ is a more malleable variable compared to student SES. If STRQ could buffer the association between low SES and student academic outcomes, this could have strong implications for interventions as improving STRQ could indirectly support academic outcomes for students with disabilities and delays living in low-income households.

Hypothesis 3d. STRQ in preschool will moderate the association between student behavior in Head Start and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student behavior ratings and kindergarten reading outcomes when strong STRQ is present. For preschool students with disabilities, behavior problems in Head Start are negatively associated with kindergarten academic outcomes (Oliver et al., 2020). Student behavior problems could also lead

to additional school-related risk factors, which can indirectly affect student academic outcomes. Students with increased behavioral concerns can be at a higher risk for disciplinary infractions and suspensions, as well as decreased engagement to complete academic activities (Kremer et al., 2016; Krezmien et al., 2006; Olivier et al., 2020). When students are not in school during a suspension or not engaged in academic tasks, academic performance can be negatively impacted. The presence of strong STRQ in Head Start could dampen the association between student behavior challenges in Head Start and kindergarten reading outcomes. Specifically, I predicted that positive STRQ would dampen the effects of student behavior difficulties in Head Start on students' kindergarten reading outcomes when student background variables were controlled as observed by a statistically significant interaction effect between student behavior and STRQ.

Hypothesis 3d, along with the hypotheses below, were supported by both theory and extant literature emphasizing the importance of STRQ for student academic outcomes (e.g., Decker et al., 2007; Hamre & Pianta, 2001; Pianta & Stuhlman, 2004). Hypothesis 3d was further supported by findings indicating that students' WJ Achievement assessment scores are more closely related to STRQ (r = .20, p <.05) compared to teacher reports of student behavior (r = -.01, p >.05) in a study of kindergarten students living in low-income settings (Mantzicopoulos et al., 2005). In another study, Pianta and Stuhlman (2004) similarly report that WJ vocabulary scores are more closely related to STRQ conflict and closeness (p < .05) compared to teacher reports of students' internalizing and externalizing behavior scores (p > .05) on the Child Behavior Checklist (CBCL). Thus, it was predicted that STRQ would moderate the association between student behavior difficulties in Head Start and kindergarten reading outcomes based on evidence that academic outcomes tend to be more closely connected to STRQ compared to student behavior challenges.

Hypothesis 3e. STRQ in preschool will moderate the association between teacher experience and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between teacher experience and kindergarten reading outcomes when strong STRQ is present. Research suggests teacher experience is connected to student academic outcomes. Specifically, higher student reading and math scores result when teachers have more years of experience (Ladd & Sorensen, 2017). Although this trend tends to level out after about ten years of experience, nonetheless students with teachers with less experience tend to demonstrate lower academic performance. Yet, the presence of strong STRQ was predicted to moderate this relationship in that STRQ in Head Start could attenuate the association between Head Start teacher experience and kindergarten reading outcomes. Specifically, I predicted that positive STRQ would dampen the effects of Head Start teacher experience on students' kindergarten reading outcomes when student background variables were controlled as observed by a statistically significant interaction effect between teacher experience and STRQ.

Similar to previous hypotheses, hypothesis 3e was supported by both theory and research emphasizing the importance of STRQ for student academic outcomes (e.g., Decker et al., 2007; Hamre & Pianta, 2001; Pianta & Stuhlman, 2004). This hypothesis was further supported by findings in the disability literature. Specifically, Zendarski and colleagues (2020) found that student reading abilities were more closely related to STRQ (r = .20, p <.01) compared to teacher experience (r = .00, p >.05) in a study with and without students with disabilities. Other academic skills, such as student vocabulary skills and math skills were also more strongly associated with STRQ (r = .16, p <.05; r = .14, p <.01), compared to teacher experience (r = -.04, p >.05; r = -.01, p >.05), respectively. Moreover, student learning skills were significantly

associated with STRQ (r = .09, p < .01), unlike the non-significant association between student learning skills and Head Start teacher years of experience (r = .20, p >.05) when analyzing the FACES 2014 data (Jeon et al., 2021). Thus, STRQ was predicted to moderate the association between Head Start teacher experience and kindergarten reading outcomes as academic outcomes tend to be more closely connected to STRQ compared to teachers' experience.

Hypothesis 3f. STRQ in preschool will moderate the association between teacher mental health and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between teacher mental health difficulties and kindergarten reading outcomes when strong STRQ is present. Initial evidence from typically developing populations indicates that teachers' mental health is connected to student outcomes. Namely, teachers who report fewer depressive-like symptoms have students with more positive wellbeing (Harding et al., 2019). I predicted that positive STRQ would dampen the effects of Head Start teacher mental health difficulties on student reading outcomes during spring of kindergarten. Particularly, teacher mental health difficulties were predicted to not be as strongly associated with student academic outcomes when STRQ is high. This would have been observed by a statistically significant interaction effect between teacher mental health and STRQ when student background variables (i.e., student cohort, student academics) were controlled.

Statistically, researchers suggest a strong association between two variables indicates a good case for including the variable as a moderator in future analyses (Memon et al., 2019). As stated above, there is strong evidence demonstrating the importance of student-teacher relationships for student academic outcomes (McGrath & Van Bergen, 2015; Pianta et al., 2008). Additional evidence supporting this hypothesis came from research. For example, academic outcomes are more closely related to STRQ (r = .20, p <.05) compared to factors contributing to

teacher stress and mental health such as difficulty of teaching assignment ($r = .18$ $p > .05$) or teaching at-risk students [($r = -.10$, $p < .05$); (Mantzicopoulos, 2005)]. This was echoed in a recent study of over 1500 Head Start students using the FACES 2014 data set. Student learning skills, such as their attention, independence, and motivation to complete academic tasks were more strongly related to STRQ in the fall ($r = .09$, $p < .01$) and spring ($r = .09$, $p < .001$) of Head Start, compared to the association between teacher depression symptoms and student learning skills in the fall ($r = -.01$, $p > .05$) or spring [($r = -.02$, $p > .05$); (Jeon et al., 2021)].

Therefore, the effect of Head Start teacher mental health on student's kindergarten reading outcomes were predicted to vary as a function of STRQ. Specifically, STRQ was predicted to attenuate the association between Head Start teacher mental health difficulties and kindergarten reading outcomes as research suggests that if teachers can form positive STRQ with students, this may be more meaningful for academic outcomes compared to their own mental health challenges. Thus, when teachers foster strong STRQ with students with disabilities, the effects of teacher mental health difficulties on student kindergarten reading outcomes were predicted to dampen.

Hypothesis 3g. STRQ in preschool will moderate the association between class size and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between class size and kindergarten reading outcomes when strong STRQ is present. For students with disabilities, class size lends to opportunities for increased academic support. Smaller class size likely improves the student-teacher ratio, allowing for additional student-teacher interactions, academic instruction, and individualized feedback (Zigmond & Kloo, 2017). Moreover, smaller class sizes have been connected to improved academic outcomes for students with disabilities as compared to their peers in larger

classrooms (Francis & Barnett, 2019). Thus, students in larger classrooms may experience poorer academic outcomes compared to their peers in smaller classrooms. It was hypothesized that positive STRQ would dampen the effects of Head Start class size on kindergarten reading outcomes as demonstrated by a statistically significant interaction effect between class size and STRQ when student background variables were controlled.

Along with research emphasizing the importance of STRQ for student academic outcomes (e.g., Decker et al., 2007; Hamre & Pianta, 2001; Pianta & Stuhlman, 2004), hypothesis 3g was supported by the notion that STRQ is more important than the frequency of interactions between students and teachers. Students in large classrooms likely have fewer interactions with their teachers due to larger student-teacher ratios. However, drawing from the parent-child literature, research suggests the quality of the relationship is more important than the quantity. For example, academic outcomes were significantly related to the quality of parent-child communication using the Early Childhood Longitudinal Study, while academic outcomes were not significantly correlated with the quantity of the parent-child communication (Zhang, 2020).

Therefore, the effect of Head Start class size on student's kindergarten reading outcomes were predicted to vary as a function of STRQ. Specifically, STRQ was predicted to attenuate the association between Head Start class size and student kindergarten reading outcomes as research suggests that if teachers can form positive STRQ with students, it may be more important than having less frequent interactions in a large class size. Moreover, strong STRQ could be more meaningful for future academic outcomes even if students have less frequent interactions with teachers when placed in a large class size. Thus, I predicted that when teachers foster strong

STRQ, the effects of Head Start class size on student kindergarten reading outcomes may be dampened.

CHAPTER 3: Method

Study Overview

The present study used data from the Family and Child Experiences Survey (FACES) 2009 to investigate the importance of STRQ for students with disabilities and delays. Dr. Kristin Rispoli and I submitted an application to the Inter-University Consortium for Political and Social Research (ICPSR) in June 2021 to request the FACES 2009 data set. ICPSR accepted the application, and the FACES 2009 data set was provided to each researcher.

My study tested one model. The model examined a) the association between STRQ and kindergarten reading outcomes for students with disabilities and delays, b) the association between the HS-STC variables and kindergarten reading outcomes for students with disabilities and delays while controlling for student background variables, and c) STRQ's moderating role between HS-STC variables and kindergarten reading outcomes while controlling for student background variables. All analyses were tested using structural equation modeling (SEM) in *Mplus* Version 8.6 (Muthén & Muthén, 1998-2021) software.

FACES Design

Data collection was completed in four waves for the FACES 2009 study, including the (1) fall semester of students' first year in Head Start (i.e. fall 2009), (2) spring semester of students' first year in Head Start (i.e., spring 2010), (3) spring semester of the three-year old students' second year in Head Start/Kindergarten year of the four-year-old students (i.e., spring 2011), and (4) spring semester of kindergarten for the three-year-old students [(i.e., spring 2012); (USDHHS, 2009a)]. My study included data from all data collection waves.

Sampling

FACES Sample

The FACES 2009 study used a multistage sampling design, including (a) Head Start Programs, (b) centers nested in programs, (c) classrooms nested in centers, and (d) children nested in classrooms (USDHHS, 2009a). The present study focused on the final nesting cluster (i.e., children nested in classrooms). The first three stages used a probability proportional to size method, followed by an equal number of samples gathered for the last stage. The rationale for this method was to lessen the unequal weighting effects on variance estimates while aiming to provide an equal probability for children to be selected for the study. Implicit and explicit stratification was used at each stage and sequential sampling was used in the first three stages. Explicit stratification involved sorting the sample into groups based on certain characteristics prior to sampling, while implicit stratification involved sorting the specified groups by certain characteristics prior to sampling. Stratification was one way the FACES 2009 study aimed to ensure the selected sample was representative of the larger sample (USDHHS, 2009a).

The sample included preschoolers, ages three through five years old, who entered Head Start for the first time. The FACES 2009 study aimed to gather data from 60 Head Start Programs, two centers from every program, and at most three classrooms from every center. In all, 60 programs, 130 centers, 486 classrooms, and 3,349 children participated in the FACES 2009 study. Based on the sample design, the large sample size should be adequate to notice meaningful differences for a variety of analyses, including those related to child outcomes according to the USDHHS(2009a). See Appendix B for the expected and actual sample size of the FACES 2009 study.

Head Start Programs. The 2007-2008 Head Start Program Information Report (PIR) database was used at the first stage of sample selection (USDHHS, 2009a). The PIR included Head Start Programs located in all 50 states and the District of Columbia, which resulted in approximately 2,600 programs. Some programs were excluded from the sample selection, such as programs that were defunded, did not directly provide services to children, located in United States territories or Puerto Rico, and under the Region XI (American Indian/Alaska Native) or Region XII (Migrant/Seasonal) regulations. In March 2009, programs were selected using stratification methods to approximate students first enrolled in Head Start. Twelve explicit sampling strata were created based on urbanicity, racial/ethnic minority enrollment, and census region. Implicit sampling strata were based on the following criteria, including (a) percentage of students with disabilities, (b) percentage of dual language learners, (c) public school district grantee status, and (d) percentage of students for which English was their primary language. Initially, researchers selected double the number of programs needed for the study in case some programs were unable to participate. Thus, programs were placed in pairs with similar explicit and implicit sampling criteria using a sequential sampling process. Then, one program from each pair was selected for the study. In the final sample of 60 Head Start Programs, 55 programs were from the initial selection and 5 programs were replacement programs (USDHHS, 2009).

Centers. The 60 Head Start Programs provided a list of expected students beginning Head Start for the first time, as well as additional sampling information in the summer of 2009. Two centers were randomly chosen from each of the 60 Head Start Programs (USDHHS, 2009). Centers were excluded if they were "partnerships" or not providing direct services. Centers were grouped based on size and geography, as centers that would likely have less than ten students enrolled in Head Start for the first time were paired with larger centers. No explicit stratification

was used. Implicit stratification was based on English Language Learner percentage. The final sample included 130 centers.

Classrooms. Head Start Centers provided information about classrooms, including (a) the number of classrooms, teachers, home visitors, and children, (b) whether the classroom operated as a full-day, morning, or afternoon program, and (c) the number of children enrolling in Head Start for the first time (USDHHS, 2009a). Classrooms with less than ten students who were new to Head Start were combined with another class. For example, the largest and smallest classrooms within the same center were combined, followed by the next largest and next smallest until all necessary classes were combined. Implicit stratification occurred based on full day or half day. Overall, the FACES 2009 study included 486 eligible classrooms.

Children. Finally, centers provided lists of students enrolling in Head Start for the first time, specifically students' names, sibling information, enrollment status, and age. The age indicator was used to assign students to a three-year-old or four-year-old cohort (i.e., Cohort 3, Cohort 4) and confirm their eligibility in the study. Students were oversampled in this wave to account for projected loss of participation. Thirty-six students enrolling in Head Start for the first time were selected from each classroom with every student having equal probability of being selected. If additional students were enrolled after the Field Enrollment Specialist site visit, they were not included in the study. One child from each family was selected, which reduced the sample to about 96.8% of the original sample. Ninety percent of parents gave consent to participate in the study (i.e., 3,349 of 3,718 children). The final sample did not include students who left Head Start before kindergarten. After considering rates of attrition, it was expected that about 55% of students would remain in the final wave of data collection. See Appendix B for additional information about the expected and final FACES 2009 study sample.

Study Sample

Consistent with the FACES 2009 study, my study's sample included three-and four-year-old students entering Head Start for the first time. Only students identified with a disability or delay during the first wave of data collection were included in the study. Student disability and delay information was gathered via the teacher child report completed in the fall of 2009. Teachers answered the following question for each student "Has any professional such as a doctor or other health or education professional mentioned this child having a developmental problem or delay, for example, any special need or disability, such as physical, emotional, language, hearing difficulty or other special need". Teachers were provided with the following responses: a) yes, b) no, or c) don't know. These data were used to determine which students were included in the present study, as students with no identified disability or delay were excluded. Of the 3,349 total children who were eligible and had parental permission for the FACES study in fall of 2009, 382 were identified as having a disability or delay according to the FACES 2009 User guide (USDHHS, 2009a).

Prior to running analyses, the data were sorted and cross examined to ensure students were in the same Head Start classroom across waves one and two, which decreased the sample from 382 students to 367 students. Of the 367 students in the present study, most students were male (61.3%) compared to female (34.9%). Students were racially diverse, with most students identifying as white (34.9%) and Hispanic/Latino (30.5%). The majority of students were above the poverty line (51.5%) and three years old (57.5%). Family incomes ranged between $0 to $5,000 (2.5%) to over $75,000 (1.6%). Most families had an annual income between $10,001 and 25,000 (51.2%). All students were identified as having a disability or delay by their teacher, specifically that a health or educational professional reported that the student had a

developmental problem or delay. Most students were identified as having a speech impairment (69.5%) and developmental delay (23.4%). Other students were identified as having a vision impairment (7.6%), blindness (0.8%), hearing impairment/hard of hearing (3.3%), deafness (0.5%), motor impairment (6.0%), mental retardation (0.3%), autism (2.7%), behavior (ADHD) (11.7%), oppositional defiant disorder (1.1%), and other (6.0%). Approximately 44% of students in the sample had an IEP or IFSP. Most students attended Head Start for four (29.7%) or five days (57.2%) a week. See Table 1 for additional demographic details and how the various categories were represented related to the variable of interest.

Table 1. Student Demographic Information with STRQ/KREAD Mean Scores

Characteristic	Participant (n=367)	Percentage	STRQ (n = 275)	KREAD (n = 206)
Gender				
Male	225	61.3	5.32 (0.53)	100.30 (16.91)
Female	128	34.9	5.32 (0.50)	106.47 (13.46)
Race/Ethnicity				
African	85	23.2	5.15 (0.57)	102.00 (19.89)
American				
Other	283	72.7	5.32 (0.49)	102.65 (14.49)
Age				
3	211	57.5	5.35 (0.48)	104.34 (15.15)
4	136	37.1	5.27 (0.57)	100.69 (16.26)
SES				

Table 1 (cont'd)

Below Poverty Threshold	157	42.8	5.38 (0.48)	104.03 (15.58)
Above Poverty Threshold	189	51.5	5.28 (0.55)	101.36 (16.25)

Most parents who completed the parent interview were mothers of the student (80.7%). Most students resided in homes that were not two parent households (45.0%), compared to students who lived in homes of parents who were married (31.3%) or not married (16.3%). Mothers' ages ranged from younger than 18 years old to over 50, with a mean age of 29.6 (SD 6.166). Fathers' ages ranged from 20 years old to over 60 years old with a mean age of 33.24 (SD 8.01). Mothers and fathers identified as white (39.8%; 36.5%), African American (23.4%; 27%), Hispanic/Latino (28.1%; 26.2%), American Indian/Alaska Native (1.1%; 0.5%), Asian/Pacific Islander (0.5%; 1.1%), Multi-Racial (1.4%; 1.9%), and an Other Race (1.4%; 1.1%), respectively. Mothers' and fathers' educational levels ranged from less than a high school diploma (26.4%; 17.2%), high school diploma/GED (34.3%; 16.9%), some college (19.1%; 7.1%), to bachelor's degree or higher (5.7%; 3.5%), respectively. Finally, family annual income ranged from below $5,000 to over $75,000, with most families endorsing an annual income between $20,000 – 25,000 a year. See Table 1.

Of the 243 teachers who participated in the study, most teachers were female (82.8%). Teachers identified racially as White (51%), African American (19.3%), American Indian/Alaska Native (2.2%), Asian/Pacific Islander (1.9%), and another race (10.6%). Most teachers had an associate degree (N = 90; 24.5%) or bachelor's degree (N = 129; 35.1%).

Experience teaching Head Start in the fall of 2009 ranged from 0 to 30 years with a mean of 8.91 years (SD = 7.16). Class sizes ranged from 11 to 20 students, with a mean of approximately 17 students per class (SD = 2.15). The teacher to child ratio in the classrooms ranged from 3.67 to 19, with a mean of 8.12 (SD = 1.98). 109 centers and 59 program were represented in this dissertation study.

Nesting

Nesting was considered as students in lower-level units were situated in Head Start classrooms, centers, and programs in higher-level units. Nesting can lead to statistical dependency, violating the assumption of independent outcomes (Moerbeek, 2004). This is problematic as standard errors may produce biased estimates as students nested within the same classrooms, centers, and programs may be more similar to each other than to students in different classrooms, centers, and programs. Thus, differences within group (i.e., students in the same classroom) and between group (i.e., students in different classrooms, centers, and programs) can be analyzed when data are nested.

Nesting was investigated by running a multi-level model to examine the intraclass correlation (ICC) in *SPSS* Version 28.0 (IBM Corp, 2021). The ICC examines how much variance the level two, three, and four cluster accounts for, with lower levels of ICC reflecting smaller differences between groups. The results of the two-level null multi-level model indicated that 98.4% of the variance in the model was at the student level, and the remaining 1.6% of the variance was at the class level.

When three-level and four-level multi-level models were completed in both *Mplus* and *SPSS*, convergence warnings were noted and the models did not converge. This is likely due to the small sample size and the fact that there are likely few students nested within the level three

(i.e., center) and level four (i.e., program). Therefore, based on the results from the preliminary analyses, I used the "Cluster" statement in *Mplus* to account for higher-level nesting effects within the data structure, as it adjusts the standard errors and chi-square fit statistic, rather than using multilevel modeling due to relatively small ICC values at the cluster level. Specifically, I clustered on the classroom variable.

Data Collection

Data for the FACES study were collected across three academic years (i.e., 2009-2012) in four waves for three-and four-year old students entering Head Start for the first time. Multiple forms of data collection occurred, including (a) child ratings completed by parents, teachers, and assessors, and direct child assessments, (b) parent surveys, (c) teacher surveys and classroom observations, and (d) leadership surveys. Data collection utilized both web-based and paper-based data collection methods, and surveys were completed via computer-assisted telephone and personal interviewing. See Table 2 for a timeline of data collection for the measures of interest.

Table 2. Data Collection by Wave

Variable	Measure	Student in Head Start		Student in Kindergarten	
		Fall 2009	Spring 2010	Spring 2011	Spring 2012
HS Reading	Assess	X			
Child Cohort	-	X			
Student Gender	PI	X			
Student Race[a]	PI	X			
Student SES	PI	X			

Table 2 (cont'd)

Student Behav.	TCR	X				
Teacher Exper.	TI	X				
Teacher MH	TI	X				
Class Size	TI	X				
K Reading[a]	Assess				X	X
STRQ	Observation		X			

Note. TCR = Teacher Child Report; PI = Parent Interview; TI = Teacher Interview; Asses = Direct Assessment; Behav. = Behavior; Exper. = Experience; MH = Mental Health.

[a]Derived Variable.

My study used student, teacher, and classroom data collected via direct student assessment, parent interview, teacher interview, teacher child report, and the STRQ data collected via a classroom observation. The first data collection wave was completed in the fall of 2009 (i.e., the student's first year in Head Start). The second wave of data collection was completed in the spring of 2010. The STRQ data collected via a classroom observation was used in my study. Finally, the third and fourth waves of data collection were completed in the spring of 2011 (i.e., four-year-old's kindergarten year) and spring of 2012 (i.e., three-year-old's kindergarten year), respectively. See Tables 3 and 4 for information regarding data collection waves and associated measure details.

Response rates were high across the waves of data collection. Ninety percent of parents provided consent to participate in the study. In fall 2009, 93% of parent interviews and 94% of student assessments were competed. 97% of teachers completed ratings for their students. In

spring 2010, 10% of students left Head Start. Of the remaining students, 86% of parent interviews and 95% of student assessments were completed. Ninety six percent of teachers completed ratings of their students. In spring 2011, 81% of student assessments were completed with kindergarten students. In spring 2012, 85% of student assessments were completed with kindergarten students.

Missing data were considered Missing at Random (MAR). MAR suggests that there is likely an underlying reason for the missing data as it is not completely random. To address these missing data from the FACES 2009 data set, the Full Information Maximum Likelihood (FIML) was used. FIML provides an estimate for each person using available data. FIML was chosen as it provides unbiased standard errors and parameter estimates when missing data are considered MAR as all available data are retained in the models (Enders & Bandalos, 2001).

Data Preparation

Among the three data files available for FACES 2009, the Child-Level data file and the Classroom/Teacher data file were utilized in my study. The Child-Level data file includes information regarding the direct child assessments, parent interviews, teacher child surveys, and classroom observations (USDHHS, 2009a). The Child-Level data file has information for all 3,349 students, even if data for children are missing at any wave. The data file is organized into three major sections that includes (1) demographic information, (2) derived or constructed variables, and (3) interview data. The Classroom/Teacher data file contains the teacher interview data completed in fall of 2009 and spring of 2010.

Data files were merged and cleaned prior to running any statistical analyses. Per the FACES 2009 user guide, Child-Level and Classroom/Teacher Level data files were merged using the CSL_ID (i.e., class ID). Data were then sorted by disability status using the variable

R1F01. All cases were retained if the student was identified with a disability or delay (n = 382). Cases were then cross-examined to ensure that students were in the same Head Start classroom in waves one and two (i.e., fall 2009 and spring 2010). Cross-examining was completed because it was important for teacher variables collected in fall 2009 to align with classroom data collected in spring 2010. Of the 382 students with an identified disability and delay, the remaining sample was 367 students across 243 classrooms.

Constructs and Measures

Multiple measures were used to answer the research questions of interest for students with disabilities and delays. The following section provides a description of the measures used in this study. See Table 3 for additional details.

Table 3. Variable Description

Description	Variable	Instrument	Time	Scale/Range	Reliability[c]
HS Reading	A1WJLWS	Assess	F09	0 - 200	0.85
Student Cohort	COHORT	-	F09	3 or 4 (cohort)	N/A
Student Gender	CHGENDER	PI	F09	Male – Female	N/A
Student Race[a]	CRACE	PI	F09	0 - African American non–Hispanic. 1 - White non–Hispanic; Hispanic/Latino; American Indian or Alaska Native; Asian or Pacific Islander;	N/A

Table 3 (cont'd)

Student SES	P1POVRTY	PI	F09	Multiracial/ biracial non–Hispanic; Other race Below poverty threshold - At or above poverty threshold	N/A
Student Behavior[a]	R1BPROB2	TCR	F09	Composite Score 0 – 15	N/A
Teacher Exper.	T1D01	TI	S10	0 - 30	N/A
Teacher MH	T1DEPCAT	TI	F09	1 - 4	N/A
Class Size	T1CSIZE	TI	F09	2 - 21	N/A
K Reading[b]	KREAD	Assess	S11 S12	0 - 200	N/A
STRQ	O2CLSSES	CLASS	S10	1 - 7	0.80

Note. TCR = Teacher Child Report; PI = Parent Interview; TI = Teacher Interview; Asses = Direct Assessment; Exper = Experience; MH = Mental Health; HS = Head Start; K = Kindergarten; F09 = Fall 2009; S10 = Spring 2010; S11 = Spring 2011; S12 = Spring 2012. [a]Only Externalizing Items in Variable Included. [b]Recoded Variable. [c]The reliability is Cronbach's Alpha.

Exogenous Study Variables

Student Gender. Student gender data were collected via parent interview in the first data collection wave, fall of 2009. The variable CHGENDER was used to measure gender. CHGENDER is a dichotomous variable in which 0 indicates female and 1 indicates male.

Student Race. Student race data were collected via parent interview in the first data collection wave, fall of 2009. The variable CRACE is a categorical variable (i.e., 1 = White, Non-Hispanic; 2 = African American, Non-Hispanic; 3 – Hispanic/Latino; 4 = American Indian or Alaska Native, Non-Hispanic; 5 = Asian, Non-Hispanic; 6 = Multiracial/Biracial, Non-Hispanic; 7 = Other Race, Non-Hispanic). Aligned with recent research (e.g., Bates & Glick, 2013; Jerome et al., 2009), CRACE was recoded into a dichotomous variable, specifically 1 indicating African American, Non-Hispanic and 0 indicating all other races.

Student SES. Student SES data were gathered via parent interview in the first data collection wave, fall of 2009. The P1POVRTY variable was used as a measure of SES, derived from parents' reports of household size and total family income of all members in the household prior to any deductions or taxes, including any money received from public assistance programs. P1POVRTY is a dichotomous variable with 0 indicating "below the poverty threshold" and 1 indicating "at or above the poverty threshold". Specifically, households with two ($14,051), three ($17,163), four ($22,025), five ($26,049), six ($ 29,456), seven ($33,529), eight ($37,220), and nine ($44,346) members were consider at or above the poverty line if their annual income was more than the amount indicated in parentheses. Conversely, if families had an annual income lower than the amount indicated in paratheses, they were considered below the poverty line.

Student Behavior. Student behavior data were collected via the teacher child report that was completed in the fall of 2009. The problem behavior questionnaire measures adverse behaviors that are linked to future school problems, such as retention and learning concerns. The scale uses questions from the shortened Personal Maturity Scale (Entwisle et al., 1987) and the Behavior Problems Index (BPI; Peterson & Zill, 1986). The BPI elicits information about both internalizing (e.g., depression, withdrawal) and externalizing (e.g., hyperactivity, aggression) behaviors. Teachers answered ten questions using a one (not true) to three scale (very true or often true) for each student. Teachers completed questions about each student's a) age-appropriate behavior, b) concentration/attention, c) fighting/hitting, d) withdrawal, e) confidence, f) nervousness, g) fidgeting, h) temper, i) sadness/depression, and j) worrying. Compared to higher scores, lower scores indicated less severe behavior or that negative behavior occurred less often.

In the FACES 2009, the variable R1BPROB2 was used as a measure of child behavior. R1BPROB2 is a total score of behavioral problem items reported by teachers with values ranging from 0 to 30. When used in the National Longitudinal Study of Youth (Berry et al., 2004), the BPI had an internal consistency of 0.89. Similarly, the internal consistency was 0.88 in the FACES 2009 study. In the present study, confirmatory factor analyses were completed for student behavior to confirm the factor structure. A summary score including only the externalizing variables (i.e., age-appropriate behavior, concentration/attention, fighting/hitting, fidgeting, and temper) was used in the final model to increase parsimony, as student problem behavior was not the main variable of interest. See the results section for additional details.

Teacher Experience. Head Start teachers reported their teaching experience during the teacher interview in the first wave of data collection, fall of 2009. Teacher experience was

measured using the T1DO1 variable. T1DO1 is a continuous variable with responses ranging from 0-30 following the teacher's answer to the following question "In total, how many years have you been teaching (including all grades and preschool)?" (USDHHS, 2009b). Higher scores indicated more teaching experience.

Teacher Mental Health. The Teacher Mental Health survey was completed via the teacher interview in the fall of 2009. The Teacher Mental Health Survey is a 12-item measure of depressive symptoms, adapted from the Center for Epidemiologic Studies of Depression Scale (CES-D; Radloff, 1997). Items include questions regarding a) feeling bothered, b) appetite, c) feeling blue, d) staying focused, e) depressed mood, f) tasks feeling effortful, g) fearfulness, h) sleep, i) frequency of talking to others, j) loneliness, k) sadness, and l) initiative. Head Start teachers rated each item based on how they were feeling in the last week using a four-point Likert scale, ranging from "rarely or never" to "most or all of the time".

The FACES 2009 study used the variable T1DEPSCO to measure teacher mental health, which is a continuous variable of the Teacher Mental Health Survey total score. Values range from 0-36. Higher scores reflect higher depression symptoms. The internal consistency was 0.80 in the FACES 2009 study. My study used T1DEPCAT as a measure of teacher mental health, which was derived using the TIDEPSCO scores. T1DEPCAT is a categorical variable that was derived from teachers' responses to the Teacher Mental Health Survey in the fall of 2009. Scores range from 1 (not depressed) to 4 (severely depressed). Specifically, Teacher Mental Health Survey scores ranging from 0-4 indicate not depressed, 5-9 indicate mildly depressed, 10-14 indicate moderately depressed, and 15 or more indicate severely depressed.

Class Size. Teachers reported their class size via the teacher interview in the fall of 2009. Class size was measured using the continuous variable T1CSIZE. Values range from 2-21 with higher numbers indicating more students enrolled in the classroom.

Classroom Assessment Scoring System (CLASS). The Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) was used as the measure of STRQ in my study. I used an observational measure, such as the CLASS, in efforts to limit measurement bias that can result when using informant-reports and to avoid teachers reporting information for both the exogenous variables (e.g., teacher experience, teacher mental health) and the moderating variable (i.e., STRQ). The CLASS has been used as a measurement tool in the FACES data series since 2006. The CLASS measures classroom quality related to both social-emotional and instructional components within the environment. The FACES 2009 included the following three CLASS domains: Classroom Organization, Instructional Support, and Emotional Support. The Classroom Organization domain measures the environmental factors such as behavior management techniques or instructional learning formats. The Instructional Support domain measures how concepts are developed in the classroom and the use of language to support instruction. The Emotional Support domain measures teacher's sensitivity to student needs, regard for student perspectives, and overall classroom climate.

In the FACES 2009 study, CLASS scores were collected via observation by a trained, third-party rater. Following a training in the fall of 2009 for the first wave of data collection, raters received an additional training specifically for CLASS observations. First, raters attended an eight-day observer training that included presentations, quizzes, examples, and practice opportunities related to the observational measures. In addition to receiving a follow-up, two-day refresher training where observational procedures were reviewed and questions were answered,

raters independently completed practice CLASS ratings to determine reliability. Each rater completed three CLASS ratings from pre-recorded videos, which were compared to the master codes by project staff members. During the last day of refresher training, in groups of three raters and one master trainer, groups observed four preschool classrooms for 20 minutes. Ten minutes after each observation were allotted for coding. Scores were compared using a reliability sheet in which raters needed to match the master trainers' codes by +/- one point for at least 80% of items to be reliable. Twenty of the twenty-two raters were reliable and deemed certified to complete the CLASS. Researchers used ongoing refresher trainings during each data collection wave to ensure reliability was maintained.

In the FACES 2009, CLASS observations lasted a minimum of four hours and occurred in the morning. CLASS observations completed in the spring of 2010 were used in my study. Observations were documented using paper forms, which were later transferred to an online format. Observers rated all CLASS components using a seven-point Likert-type scale ranging from one "minimally characteristic" to seven "highly characteristic". Higher scores indicated a more positive STRQ.

My study only used CLASS data related to the Emotional Support domain to measure STRQ due to its alignment with the definition and conceptual understanding of STRQ in the literature (Walker & Graham, 2021). In the FACES 2009 study, Emotional Support was measured using the O2CLSSES variable. The O2CLSSES variable is a measure of the CLASS Emotional Support domain, which was created as a mean score of the Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspective dimensions. Prior to creating a mean score, Negative climate scores were reverse coded. In the FACES 2009 study, the internal consistency for Emotional Support was 0.91 and average inter-rater reliability was 87%

(USDHHS, 2009a). Correlations between two raters on the Emotional Support Domain and the a) full Early Childhood Environmental Rating Scale -Revised (ECERS-R; r = .68) and the b) shortened ECERS-R (r =.76) were moderate. The ECERS-R measures aspects of the classroom quality, such as the language, interactions, activities, routines, and structure embedded in the program.

The CLASS is a reliable and valid measure. In a study of 224 preschool classrooms, participants' ratings of dimensions within each domain were strongly associated (La Paro et al., 2004). For example, within the Emotional Support domain, the dimension of Positive Climate was negatively associated with Negative Climate (r = -.64, p < .01) and positively associated with Teacher Sensitivity (r = .82, p < .01). Negative Climate was negatively associated with Teacher Sensitivity (r = -.54, p < .01). The correlation between Positive Climate and Teacher Sensitivity was the strongest among the dimensions, which points to the connection between teacher practices and the classroom environment when gauging STRQ. Statistically significant correlations between dimensions were also reported for the other two CLASS domains (La Paro et al., 2004). In the same study, the CLASS was found to significantly correlate with another commonly used measure of classroom quality, the Early Childhood Environmental Rating Scale (ECERS). Specifically, the Emotional Climate domain was most strongly correlated with the overall ECERS score (r = .52, p <. 001), student-teacher interactions score (r = .58, p < .001), and language reasoning score (r = .47, p < .001), suggesting that the CLASS similarly measures these constructs. High reliability scores are also reported from publishers of large-scale secondary data sets. When used in the Family and Child Experiences Survey (FACES) 2014 study, the internal consistency for the CLASS Emotional Support domain was 0.91 and the

average inter-rater reliability was 87% (Pianta et al., 2008). Thus, multiple findings allude to the reliability and validity of the CLASS, especially when used with early childhood populations.

In lieu of using the Emotional Support variable from the FACES 2009 study, a latent term constructed from the four dimensions of the Emotional Support domain was used as a measure for STRQ. Rather than using the derived variable for Emotional Support, this approach was chosen to increase precision in the model, especially with the construct of interest: STRQ. A confirmatory factor analysis (CFA) was completed prior to including the latent term in the primary analysis to examine if the factor structure for STRQ held for students with disabilities and delays. See Appendix C and the results section for additional information.

Endogenous Variable

Woodcock-Johnson III Tests of Achievement–Third Edition (WJ-III). The Woodcock-Johnson III Tests of Achievement-Third Edition (WJ-III) Letter-word Identification was used as a reading measure in my study. The Letter-word Identification subtest is a direct assessment that requires students to identify letters and words from a test book. In the FACES 2009 study, the WJ-III was administered by trained assessors. Assessors received field staff training on the WJ-III, passed a standard certification (>90%), and practiced assessment administration for three hours each week leading up to data collection (USDHHS, 2009a). If assessors did not pass the standard certification, they submitted a video recording of an assessment for review and had a quality assurance staff member observe their first assessment during the FACES data collection.

WJ-III standard scores were used to estimate student reading performance. Although W-scores/Item Response Theory (IRT) scores reflect changes in absolute functioning across waves, the FACES 2009 User guide reports that the fall 2009 IRT scores for letter-word performance

were calibrated with data in later waves due to insufficient sample size in the first wave. To avoid using derived variables based on data from multiple waves, standard scores were used in my dissertation study.

The WJ Letter-word Identification subtest administered to students in the fall of 2009 was used as an autoregressive control. Specifically, the variable A1WJLWS, the fall 2009 WJ Letter-Word Identification standard score, was included as a control variable. When reviewing other studies, using an autoregressive control is a common practice (e.g., Allen et al., 2017) to understand the phenomena of interest, which in this case is kindergarten reading outcomes. Studies indicate strong effects between prior and current academic performance (Croninger et al., 2007; Ramsook et al., 2020), which also supported my decision to include this variable in the study as students' performance on the WJ Letter-word Identification subtest in Head Start is likely correlated to their performance on the subtest in kindergarten. For wave one of data collection, the reliability for the WJ Letter-word Identification subtest was $\alpha = 0.85$ (USDHHS, 2009a).

The WJ Letter-word Identification subtest completed by students in their kindergarten year was used as an endogenous variable for all research questions. During students' kindergarten year, direct assessments were completed by Cohort 3 (i.e., spring 2012, wave 4) and Cohort 4 (i.e., spring 2011, wave 3). For wave three and four of data collection, reliability for the WJ Letter-word Identification subtest was $\alpha = 0.93$ (USDHHS, 2009a). Specifically, variables A3WJLWS and A4WJLWS were used as a measure of reading performance. As recommended by the FACES 2009 User guide (USDHHS, 2009a), the kindergarten variable (KGYEAR) was used to determine from which wave to select data for the outcome variable. A variable (i.e.,

KREAD) was created merging data from both cohorts using scores from students' kindergarten year.

Exogenous Control Variable

Student Cohort. The variable COHORT was used as a measure of student cohort upon entering Head Start in fall of 2009, either Cohort 3 or 4. Cohort 3 included three-year old students and Cohort 4 included four-year old students. This was used to control for potential maturation effects and differences in time enrolled in Head Start, as students in Cohort 3 spent one extra year in the program before entering kindergarten.

Variables

For this study, it was important to distinguish the observed variables from the latent variables. An observed variable has been collected in the study and a latent variable is used to measure a hypothetical construct (Kline, 2015). In my study, the following variables were considered observed variables: (a) Student Cohort, (b) Student Gender, (c) Student Race, (d) Student SES, (e) Teacher Experience, (f) Teacher Mental Health, (g) Class Size, and (h) Student Reading Scores. There were two proposed latent terms for my study: Student Behavior and STRQ. After completing CFAs, the Student Behavior variable was included as an observed variable (i.e., a summary score of externalizing behaviors) to increase parsimony in the model. STRQ, the variable of interest in my study, was kept as a latent term in the primary analysis. See Table 3 for variables and measure details.

Statistical Analyses

The following sections outline the data analytic plan for the descriptive, preliminary, and primary analyses. Data analyses were completed using *Mplus* Version 8.6 (Muthén & Muthén, 1998-2021).

Descriptive Analyses

Frequency, descriptive, and intercorrelation analyses were completed prior to running preliminary and primary analyses to further understand the variables, which helped inform which variables were appropriate to include in the study. See the results section for further details.

Preliminary Analyses

Confirmatory factor analyses (CFA) were conducted prior to running the final SEM model to understand the factor structure for students with disabilities and delays. Specifically, separate CFAs were conducted for Student Behavior and STRQ. After completing the CFAs, a summary score of externalizing behaviors was used for the Student Behavior variable to create a more parsimonious model for the primary analysis. However, STRQ is the term of interest in the present study, so the latent structure was retained for the primary analysis.

Model fit was examined and reported for all CFAs using multiple fit indices as recommended by research (Hu & Bentler, 1999). Including multiple fit indices is important because the Chi-square test uses a binary hypothesis to estimate model fit (i.e., $p \geq .05$ supports model fit), but this alone is not sufficient to indicate good model fit (Kline, 2015). Thus, the Chi-square test (Bentler & Bonnet, 1999), Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Hu & Bentler, 1999), Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993), and Standardized Root Mean Square Residual (SRMR; Hu & Bentler, 1998) were examined. The Chi-square test and SRMR are measures of absolute fit, examining the misfit that is present in the current model compared to a fully saturated model. The CFI and RMSEA are both non-centrality based fit indices. The non-centrality-based indices estimates fit using an alternative hypothesis model among the sample population (Kline, 2015). Finally, the TLI is a relative fit index. A relative fit index compares the relative improvement of this model

compared to the null model (Kline, 2015). For this study, good model fit was concluded if the Chi-square test is not significant (Bentler & Bonnet, 1980), CFI and TLI are close to 0.95 (Hu & Bentler, 1999), RMSEA values are near 0.05 and SRMR values are less than 0.06 (Hu & Bentler, 1999).

Factor loadings were also examined for significance and magnitude, with 0.30 or 0.40 considered meaningful (Hair et al., 1998). Prior to conducting the CFAs, data were prepared examining for outliers, normality assumptions, and missing data. All CFAs imposed identification constraints, specifically factor variances were fixed to one. This method was chosen to standardize the CFA outcome for interpretation as the indicator scale is arbitrary in these analyses. Using Jackson et al., (2009)'s recommendations for CFA reporting, multiple parameter estimates were presented in a table format following the analyses, such as (un)standardized coefficients, standard errors, p-values, and r squared values.

Primary Analysis

The primary analysis used structural equation modeling (SEM). According to Kline (2015), there are six main steps to conducting an SEM, including (1) Specifying the Model, (2) Identifying the Model, (3) Choosing Measures and Collecting Data, (4) Checking Model Fit, Interpreting Estimates, and Considering Other Models, (5) Respecifying the Model, and (6) Reporting the Results. The following sections use this framework to outline the data analytic plan.

Research question one examined the association between the latent variable, STRQ, and kindergarten reading outcomes for students with disabilities and delays. Research question two examined the association between HS-STC variables and kindergarten reading outcomes for students with disabilities and delays, while controlling for student background variables.

Research question three examined the moderating role of STRQ between the association between HS-STC variables and kindergarten reading outcomes for students with disabilities and delays, while controlling for student background variables. All research questions were tested using the same model. The model parameter estimates and interaction effects were examined.

Step 1: Specifying the Model. Model specification was the first step in SEM. The model included nine exogenous variables (i.e., student cohort, Head Start reading scores, student gender, student race, student SES, student behavior, teacher experience, teacher mental health, and class size) regressed on one endogenous variable (i.e., kindergarten reading scores). The endogenous variable had a disturbance, or error term. All exogeneous variables also had individual variances. Covariances were included between the exogenous variables on the same level (i.e., student background variables; HS-STC variables).

A moderating variable, STRQ, was also included in the model. STRQ was a latent term with four indicators (i.e., Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives). See Appendix D for the conceptual path diagram for the primary model.

Step 2: Identifying the Model. Step two included identifying the model. Relevant concepts related to model identification include the number of freely estimated parameters, amount of information, and the degrees of freedom. The amount of information was calculated using the following formula $(p\ (p + 1))/2$, where p was equal to the number of observed variables in the model. The degrees of freedom were calculated using the following formula: $(p\ (p + 1)/2 - q$, where p was the number of observed variables in the model and q was the number of estimates in the model. An over-identified model, which is required for the analysis, occurs when the

degrees of freedom are greater than or equal to zero or are greater than or equal to the number of parameters in the model. This requirement also satisfies the counting rule (Kline, 2015).

The model included 14 observed variables. There were nine exogeneous variables, four indicators loaded onto the latent factor, and one endogenous variable. Like the CFA that was completed in the preliminary analyses, the mean structure variance was fixed to one for identification. Using (14(14 + 1)/2), the number of observations equaled 105. There were ten direct effects from the exogeneous variables to the endogenous variable, seven direct effects related to the moderating variable, and four factor loadings. There were nine variances on the exogenous variables, four error terms on the indicators, and one disturbance on the endogenous variable. There were 36 covariances between exogenous variables in the model. There were 21 covariances between moderating variables. Thus, the number of estimated parameters equaled 92 (i.e., 10 + 7 + 4 + 9 + 4 + 1 + 36 + 21 = 92). Thus, the degrees of freedom for this model were 13 (i.e., (14(14 + 1)/2 – 92) = 13). This model was overidentified, satisfying the counting rule presented by Kline (2015), allowing for model fit and parameters to be estimated.

Step 3: Choosing Measures and Collecting Data. The measures and data collection for the study were discussed in the above sections. Refer to Tables 3 and 4 for an overview of the data collection schedule and variable descriptions.

Step 4: Checking Model Fit, Interpreting Estimates, and Considering Other Models. The maximum likelihood estimation was used to estimate the parameters in the model. The maximum likelihood estimation assumes normality for continuous endogenous variables (i.e., kindergarten reading outcomes). After running descriptive analyses, normality assumptions were examined and were not violated for kindergarten reading outcomes. To address missing data, full information maximum likelihood (FIML) was used. All continuous exogenous variables and the

endogenous variable were standardized due to differences in scaling (Asparouhov & Muthén, 2021). In the final model, the mean of STRQ was fixed to zero for ease of interpretation of the findings. Finally, grand mean centering was used for continuous exogenous variables. Centering did not change the values of the interaction term or R squared but was helpful for the variables in which zero was not included in the distribution and helped avoid multicollinearity issues (Kline, 2015). Multicollinearity is problematic as it hinders the model precision. When exogenous variables are strongly correlated, they are unable to predict the endogenous variable independently.

Absolute fit indices were not available due to the use of a latent interaction in the primary model. The Akaike Information Criteria (AIC), the Bayesian Information Criterion (BIC), and the Log Likelihood Index were produced. In general, smaller AIC and BIC values indicate a better fitting model (Kenny, 2020). The AIC changes with model complexity, and the BIC is closely tied to sample size (Kline, 2015). Comparative fit indices can only compare nested models. For nested models, the Likelihood ratio test is the difference in log-likelihood between the simplest model and the model with additional parameters (-2LL). For this dissertation study, no nested or comparative models were examined as they were outside of the scope of this particular study.

Both standardized and unstandardized results were reported. The estimate between STRQ and kindergarten reading outcomes were reported to answer research question one. Parameter estimates, standard errors, and p-values from the regressions between HS-STC variables and kindergarten reading were reported to answer research question two. Finally, the interaction effects between the moderator and the exogenous variables were also reported to answer research question three. Detailed interpretations of the results are provided below. Finally, for the

significant interaction, an interaction plot was completed to visually examine the interaction effect.

Step 5: Respecifying the Model. The primary model converged, however; aligned with recommendations from Kline (2015), the model would have been respecified and other equivalent models would have been examined if the model did not converge.

Step 6: Reporting the Results. Finally, model results were reported using details about the SEM and the analytic process. Direct parameter estimates, interaction effects, and covariances were reported. Both standardized and unstandardized results were reported, along with p-values. Many of the statistical results are presented in tables and supplemented with a narrative explanation of the findings. Additionally, a plot was provided for the significant interaction effect. Supplementary analyses were also completed to further understand the signficant moderating effect. See the results section below.

Ethical Considerations

Prior to beginning the study, an application was completed and submitted to Michigan State University's Human Research Protection Program (HRPP) for determination of non-human subjects research. The application included details of the study, as outlined above. Once completed, the application was uploaded to Michigan State University's Click Research Compliance System. The project began following HRPP designation of it as non-human subjects research (see Appendix E for this determination letter) and thus not in need of HRPP review or approval. Any questions or comments from the IRB were addressed throughout the duration of the study. Specifically, the IRB was provided with the data use agreement per request and clarification was provided regarding the FACES 2009 data set. The IRB was notified that 1. the FERPA exemption did not apply to the FACES 2009 data set as data were not taken from

students' educational record, and 2. ICPSR does not hold any personal identifiers associated with the data set. Data were reported honestly regardless of the findings.

Ethical considerations were also relevant for the FACES 2009 data. Security safeguards are in place to protect the data and identity of FACES participants. While applying for the FACES 2009 data, researchers sign a user agreement, agreeing to the following terms and conditions: 1. Data are to be used for scientific research and teaching (non-commercial); 2. Researchers should not try to identify participants; 3. Acknowledgement of ICPSR, funders, and depositors are given during publication; 4. ICPSR should receive two copies of the publication; 5. Researchers should not copyright data; 6. Data should not be shared with others and should be stored securely; 7. ICPSR should be notified of data error; and 8. No legal responsibility can be placed on ICPSR. ICPSR also reviews a summary of the proposed research project before releasing data. For my study, the user agreement was signed and ICPSR approved the summary statement.

For my dissertation study, I stored the FACES 2009 data on my password protected computer. I purchased the *Mplus* (Muthén & Muthén, 1998-2017) statistical software, rather than using software located on a Virtual Desktop, to keep the data protected and off an online platform. Data were not shared with anyone. The identification of participants was not sought. In addition to these safeguards, the FACES 2009 study protects the identity of participants by using unique IDs for all individuals who provided data for the study. Classrooms are also given unique IDs to protect confidentiality. The unique IDs are over five characters long to protect participant information. Thus, ethical implications were considered throughout the study in terms of IRB approval and data management.

CHAPTER 4: Results

Descriptive Analyses

Frequency

First, frequency data were completed for dichotomous variables included in the study. The dichotomous variables in the study included student cohort, student gender, student race, and student SES. The frequency analyses indicated that most students were in the three-year-old cohort (56.1%), male (61.3%), and identified as a race other than African American (72.8%). A similar number of students were above (51.5%) and below (42.8%) the poverty threshold. Regarding teachers' mental health, most teachers reported not being depressed (61.3%). See Table 4 for additional details.

Table 4. Categorical Variable Frequency Table

Variable Description	Variable Name	N	Percentage
Student Cohort	COHORT		
3-year-old		206	56.1
4-year-old		161	43.9
Student Gender	CHGENDER		
Female (0)		128	34.9
Male (1)		225	61.3
Missing		14	3.8
Student Race*	CRACE		
Other Race (0)		267	72.8
African American (1)		85	23.2
Missing		15	4.1

Table 4 (cont'd)

Student SES	P1POVRTY		
Below the Poverty Threshold (0)		157	42.8
Above the Poverty Threshold (1)		189	51.5
Missing		21	5.7
Teacher Mental Health	T1DEPCAT		
Not Depressed		225	61.3
Mildly Depressed		100	27.2
Moderately Depressed		23	6.3
Severely Depressed		14	3.9
Missing		5	1.4

Note. *Recoded variable. N = 367.

Descriptive

Descriptive statistics for all variables were completed to examine variable characteristics and to test the normality of the data. Specifically, the mean, standard deviation, skewness, kurtosis, and a histogram with a normal curve overlay were examined for variables. When examining the mean and standard deviations, continuous variables indicated moderate variability or dispersion of the data. See Table 5. Data are suggested to be normally distributed if skewness values range between +2 and -2 and kurtosis values range from -3 to +3 (Garson, 2012). Normality assumptions were not violated for skewness and kurtosis for all variables. In the visual analysis of the histograms, all data were in the shape of a normal bell curve, also

supporting the normality assumption. Finally, although a latent term was used in the primary analysis for STRQ, the FACES 2009 STRQ variable (i.e., O2CLSSES) was examined to learn more about the construct. The O2CLSSES variable ranged from 2.50 to 6.38 and had relatively little variability (M = 5.35, SD = 0.52).

Table 5. Variable Descriptive Statistics

Variable Description	Valid N	Min	Max	*M* (SD)	Skewness	Kurtosis
Student Cohort	367	3	4	3.44 (0.49)	0.248	-1.939
Student Gender	353	0	1	0.64 (0.48)	-0.572	-1.673
Student Race	352	0	1	0.24 (0.43)	1.213	-0.531
Student SES	346	0	1	0.55 (0.50)	-0.186	-1.965
Student Behavior	367	5	15	8.67 (2.78)	0.408	-0.864
HS Reading	238	67	161	93.95 (18.41)	0.788	0.641
Teacher Experience	308	0	30	13.28 (8.50)	0.458	-0.874
Teacher Mental Health	362	1	4	1.52 (0.78)	1.565	1.982
Class Size	361	11	20	16.77 (2.18)	-0.494	0.350
K Reading	206	51	164	102.57 (15.77)	-0.312	1.358
STRQ	275	2.50	6.38	5.35 (0.52)	-1.085	2.775

Note. Gender (0 = Female; 1 = Male); Race (0 = Other Races; 1 = African American); Child SES (0 = Below the Poverty Threshold; 1 = Above the Poverty Threshold); HS = Head Start; K = Kindergarten; Total Sample N = 367. STRQ scores can range from 1-7.

Inter-Correlations and Covariances

Next, correlations between variables were completed to see if there were significant correlations prior to running the primary analysis. The strongest correlations were between teacher experience and kindergarten reading outcomes ($r = .283$, $p < .000$), cohort and class size ($r = .233$, $p < .000$), and STRQ and teacher experience ($r = .276$, $p < .000$). Although significantly related to one another, the correlations were rather weak. None of the correlations resulted in extreme multivariate collinearity as all correlations were less than 0.90 (Kline, 2015). See Table 6 for the correlation matrix for all observed study variables.

The covariance between variables of the primary model were also examined. Most variance was shared between class size and the latent interaction between teacher mental health and STRQ. The covariance suggests that these variables tend to vary together negatively, indicating that as one variable is positive the other tends to be negative. Specifically, when Head Start students are in a larger class, the interaction between STRQ and teacher mental health tends to be negative. The covariances between variables should be considered when interpreting the final model results as these variables tend to move together. See Table 7 for the primary model's full covariance matrix for latent variables.

Table 6. Correlation Matrix

Var.		1	2	3	4	5	6	7	8	9	10	11
1.	ρ	1.000										
	p	-										
2.	ρ	-.076	1.000									
	p	.148	-									
3.	ρ	.094	.040	1.000								
	p	.077	.453	-								

Table 6 (cont'd)

4.	ρ	.063	.140**	-.025	1.000							
	p	.231	.007	.645	-							
5.	ρ	-.038	.220**	.085	-.075	1.000						
	p	.472	.000	.112	.153	-						
6.	ρ	.087	-.089	-.088	.001	-.229**	1.000					
	p	.097	.087	.098	.988	.000	-					
7.	ρ	-.004	.023	.032	.038	.031	-.053	1.000				
	p	.941	.665	.556	.473	.551	.314	-				
8.	ρ	.050	-.047	-.014	.078	-.003	.010	.034	1.000			
	p	.343	.367	.797	.137	.961	.843	.511	-			
9.	ρ	.233**	.022	.062	.045	-.038	.046	.076	.061	1.000		
	p	.000	.668	.250	.390	.467	.381	.144	.241	-		
10.	ρ	.057	-.057	-.023	.041	-.148**	.114*	.283**	.030	.132*	1.000	
	p	.278	.272	.674	.429	.004	.030	.000	.562	.011	-	
11.	ρ	-.092	.000	-.052	-.032	-.050	.021	.276**	.023	.010	.210**	1.000
	p	.080	.997	.332	.537	.336	.683	.000	.663	.846	.000	-

Note. Var. = Variable; 1 = Student Cohort; 2 = Student Gender; 3 = Student Race; 4 = Student SES; 5 = Student Behavior; 6 = Head Start Reading; 7 = Teacher Experience; 8 = Teacher Mental Health; 9 = Class Size; 10 = Kindergarten Reading; 11 = Student-Teacher Relationship Quality.

** = Correlation is significant at the .001 level (two-tailed)

*= Correlation is significant at the .05 level (two-tailed)

Table 7. Covariance Matrix for the Primary Model Latent Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.0																	
2	-.01	1.0																
3	.02	.01	1.0															
4	.01	.00	-.00	1.0														
5	-.01	.11	.03	-.05	1.0													
6	-.05	.03	.01	-.01	-.18	1.0												
7	.05	.02	.02	.01	.13	-.01	1.0											
8	.02	-.02	.00	.03	.03	-.05	.02	1.0										
9	.10	.00	-.04	-.01	-.00	.03	.05	-.01	1.0									
10	-.03	.01	-.11	-.05	.01	.02	-.02	-.18	.10	1.0								
11	-.02	.00	-.07	-.03	.00	.01	-.01	.06	-.11	.64	.64							
12	-.01	.00	-.03	-.01	.00	.00	.00	.02	-.04	.24	.16	.25						
13	-.02	.00	-.06	-.03	.00	.01	-.01	.05	-.10	.55	.35	.14	.55					
14	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.11	.04	-.05	1.0				
15	.00	.00	-.00	-.00	.00	.00	.00	.00	-.00	.01	.03	.03	.02	.13	1.0			
16	-.05	.01	-.17	-.08	.01	.03	-.03	.14	-.27	1.5	.95	.36	.85	.03	.03	2.9		
17	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	-.00	.06	.00	-.00	.06	-.02	1.0	
18	-.06	-.07	-.01	-.05	-.04	.44	-.16	.05	-.01	.08	.11	-.00	.03	-.05	.07	.17	-.04	1.0

Note. Two-tailed test; 1 = Student Cohort; 2 = Student Gender; 3 = Student Race; 4 = Student SES; 5 = Student Behavior; 6 = Head Start Reading; 7 = Teacher Experience; 8 = Teacher Mental Health; 9 = Class Size; 10 = Student-Teacher Relationship Quality (STRQ); 11 = GenderxSTRQ; 12 = RacexSTRQ; 13 = SESxSTRQ; 14 = ProblemBehaviorxSTRQ; 15 = TeacherExperiencexSTRQ; 16 = TeacherMentalHealthxSTRQ; 17 = ClassSizeXSTRQ; 18 = Kindergarten Reading.

Preliminary Analyses

Student Behavior

Prior to running the primary analysis, a series of confirmatory factor analyses (CFA) were completed for the Student Behavior variable. First, a CFA was completed for the Student Behavior items that were included in the FACES 2009 study to confirm the measurement of the student behavior construct for students with disabilities and delays. A one-factor solution indicated relatively poor fit. See Table 8 for the CFA model fit. See Table 9 for the CFA factor loadings and R-Squared results.

Table 8. Student Behavior Confirmatory Analyses Model Fit

	1. One-Factor CFA	2. Two-Factor CFA	3. One-Factor CFA*
Chi Square	373.981	217.881	53.568
	(p = 0.00)	(p = 0.00)	(p = 0.00)
CFI	0.838	0.912	0.970
TLI	0.792	0.884	0.939
RMSEA	0.162	0.121	0.163
SRMR	0.134	0.099	0.063

Note. CFA = Confirmatory Factor Analysis.

*Only Externalizing Items.

Table 9. Student Behavior One-Factor CFA Unstandardized Factor Loadings and R-Squared

Observed Variable	Estimate	S.E.	R-Squared	S.E.
Age-Appropriate Behavior	0.430*	.039	0.331*	.047
Difficulty Concentrating	0.605*	.038	0.601*	.046
Fighting/Hitting Others	0.403*	.036	0.337*	.047
Withdrawal; Keeps to Self	0.118*	.040	0.028	.019
Low Confidence	0.239*	.037	0.132*	.037
Nervousness	0.311*	.032	0.264*	.046
Fidgeting; Difficulty Sitting Still	0.639*	.038	0.637*	.044
Temper Tantrums	0.407*	.037	0.334*	.048
Sadness; Depression	0.098*	.027	0.043	.023
Worrying	0.079*	.024	0.035	.021

Note. *Significant at $p < .05$.

Due to the poor model fit of the one-factor CFA, a two-factor CFA was completed with the data. I made this decision after examining the Student Behavior items and referencing evidence from the field. After reviewing the items that were included in the FACES 2009 Student Behavior summary score, items could be interpreted as either externalizing (e.g., fidgeting, fighting) or internalizing behaviors (e.g., withdrawal, depressed) based on common psychological interpretations of behavior. Model fit improved for the two-factor CFA. See Table 8. As predicted, the results from the two-factor CFA suggested that five items significantly loaded onto an externalizing factor [i.e., a) age-appropriate behavior, b) difficulty concentrating, c) fighting/hitting others, d) fidgeting; difficulty sitting still, and e) temper tantrums] and five items significantly loaded onto an internalizing factor [i.e., a) withdrawal; keeps to self, b) low

confidence, c) nervousness, d) sadness; depression, and e) worrying]. The results aligned with my hypothesis. Higher factor loadings resulted for the externalizing variables compared to the internalizing variables. All factor loadings were significant at $p < .001$. See Table 10 for the two-factor CFA results.

Table 10. Student Behavior Two-Factor CFA Unstandardized Estimates and R-Squared

Observed Variable	Externalizing	Internalizing	R-Squared
Age-Appropriate Behavior	0.658*		0.432
Difficulty Concentrating	0.875*		0.776
Fighting/Hitting Others	0.698*		0.487
Fidgeting; Difficulty Sitting Still	0.884*		0.781
Temper Tantrums	0.740*		0.548
Withdrawal; Keeps to Self		0.490*	0.240
Low Confidence		0.634*	0.402
Nervousness		0.820*	0.673
Sadness; Depression		0.669*	0.447
Worrying		0.661*	0.437

Note. *Significant at $p < .05$.

In lieu of having two summary scores for behavior in the final model, with one summary score for internalizing behaviors and one summary score for externalizing behaviors, a final CFA was completed including only externalizing items. In addition to the strong, significant factor loadings for the externalizing items produced from the two-factor CFA, externalizing behaviors were chosen for my study based on research that indicates that parent and teacher reports of

externalizing behavior exhibited by students with disabilities are predictive of students' reading achievement (Richards et al., 1995). In addition, externalizing behaviors are a risk factor for a lack of engagement in the classroom for young students (Oliver et al., 2020), which can negatively affect academic outcomes (Lei et al., 2018).

The CFA model with only externalizing items resulted in acceptable model fit, such that x^2 = 53.568 was significant at 0.000, TLI = 0.939, CFI = 0.970, RMSEA = 0.163, and SRMR = 0.063. See Table 8. While relatively poor fit is suggested by the RMSEA = 0.163 and Chi square (x^2 = 53.568) significant at p = 0.000, significant chi-squares can often result with moderately large sample sizes (Bentler & Bonett, 1980), such as in this dissertation study. The RMSEA summarizes the errors of approximation by taking the model complexity into account (Browne & Cudeck, 1993). A poor RMSEA suggests that there is a large amount of error per degree of freedom in the model. However, like the Chi square, a poor RMSEA fit may be attributed to the RMSEA's sensitivity to sample size, model complexity, and models with small degrees of freedom (Browne & Cudeck, 1993), such as in this CFA in which the degrees of freedom were equal to five.

The CFA resulted in factor loadings that were all statistically significant. The standardized values ranged from 0.637 to 0.893. The r-squared values, which examine how much an indicator loads onto the factor, suggested that fidgeting (R^2 = 0.797, p = .000) and concentration/attention (R^2 = 0.772, p = .000) most strongly account for the Student Behavior factor. See Table 11 for additional information. See Appendix F for the path diagram with standardized results. Based on the CFA results, a summary score for externalizing student behavior was included in the final model.

Table 11. Student Behavior Externalizing Behaviors One-Factor CFA Factor Loadings and R-Squared

Observed Variable	Estimate	S.E.	R-Squared	S.E.
Age-Appropriate Behavior	0.637*	.040	0.406*	0.052
Difficulty Concentrating	0.879*	.026	0.702*	0.045
Fighting/Hitting Others	0.711*	.038	0.505*	0.055
Nervousness	0.893*	.021	0.797*	0.037
Fidgeting; Difficulty Sitting Still	0.718*	.038	0.515*	0.055

Note. *Significant at p <.05.

STRQ

Prior to running the STRQ CFA, the intercorrelations and covariances between the CLASS Emotional Support Domain indictors (i.e., Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives) were examined. The standardized correlations of the STRQ indicators suggest that Teacher Sensitivity was most strongly correlated with Regard for Student Perspectives (r = .746, p < .000). All correlations were significant, suggesting that the indicators all measure a similar construct (i.e., STRQ). See Table 12 for the correlation matrix for the STRQ indicators. Similarly, when the covariances between STRQ indicators were examined, the most variance was shared between Teacher Sensitivity and Regard for Student Perspectives (b = 0.334, p < .001). Thus, these items tended to move together such that teachers with high teacher sensitivity tended to also have positive regard for student perspectives. See Table 13 for the covariance matrix of STRQ indicators.

Table 12. STRQ Indicator Correlation Matrix

Variable		1	2	3	4
1. Positive Climate	ρ	1.000			
	p	-			
2. Negative Climate	ρ	-0.517**	1.000		
	p	.000	-		
3. Teacher Sensitivity	ρ	0.638**	-0.479**	1.000	
	p	.000	.000	-	
4. Regard for Student Perspectives	ρ	0.500**	-0.380**	0.746**	1.000
	p	.000	.000	.000	-

Note. **Significant at $p < .01$.

Table 13. STRQ Indicator Covariance Matrix

Variable		1	2	3	4
1. Positive Climate	ρ	1.000			
	p	-			
2. Negative Climate	ρ	-0.181**	1.000		
	p	.000	-		
3. Teacher Sensitivity	ρ	0.288**	-0.159**	1.000	
	p	.000	.000	-	
4. Regard for Student Perspectives	ρ	0.236**	-0.132**	0.334**	1.000
	p	.000	.000	.000	-

Note. **Significant at $p < .01$.

Next, a one-factor CFA was completed for STRQ using the four items included in the CLASS Emotional Support Domain (i.e., Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives). The model fit indices indicated adequate to moderate fit when examining the TLI = 0.873, CFI = 0.958, and SRMR = 0.044. Like the Student Behavior CFA, poor fit was suggested by the RMSEA = 0.189 and Chi square (x^2 = 21.659) significant at p = 0.000. However, significant chi-squares can often result with moderately large sample sizes (Bentler & Bonett, 1980), such as in the current study. A poor RMSEA fit may be attributed to the RMSEA's sensitivity to sample size and to models with small degrees of freedom (Browne & Cudeck, 1993), which is meaningful for this CFA due to the small degrees of freedom (i.e., df = 2).

I decided to retain the STRQ model rather than consider changes to improve model fit for multiple reasons. First, the factor structure used in my study is supported by extant research. The authors of the CLASS, as well as independent research teams have suggested that the CLASS Emotional Support domain is comprised of four dimensions (Li et al., 2020; Pianta et al., 2008). Second, because the CLASS Emotional Support domain measured the major construct of interest (i.e., STRQ), I wanted to keep the factor structure consistent with how researchers and clinicians commonly understand the measure. The CLASS has been used in multiple studies to understand student-teacher relationships (e.g., Hamre et al., 2013; La Paro et al., 2004), and when comparing the results of my dissertation study to other published studies, having a mutual understanding of STRQ was important. I also retained the proposed factor structure as the observational measurement of STRQ was a key consideration in my dissertation as researchers frequently use informant-reports to measure STRQ (e.g., Blacher et al., 2014). Finally, I used a four-factor

model for STRQ based on practical implications. When completing the CLASS, clinicians administer and interpret the CLASS Emotional Support domain using all four dimensions, rather than eliminating aspects of the measure. Thus, along with the adequate model fit, both research and practical implications influenced my decision to retain the proposed factor structure for the CLASS Emotional Support domain.

The results of the STRQ CFA indicated that the standardized factor loadings were all statistically significant. The Negative Climate indicator loaded negatively on the factor (b = -0.527) and the remaining three indicators (i.e., Positive Climate, Teacher Sensitivity, Regard for Student Perspectives) loaded positively onto the factor with standardized estimates ranging from b = 0.684 - 0.940. See Table 14 for the STRQ CFA standardized and unstandardized results. In addition, the r-squared values were all statistically significant and suggest that Teacher Sensitivity (R^2 = 0.884, p = .000) most strongly accounted for the STRQ factor. See Table 15 for additional details. See Appendix G for a path diagram with standardized results.

Table 14. STRQ Unstandardized and Standardized CFA Factor Loadings

	Unstandardized		Standardized		
Observed Variable	Estimate	S.E.	Estimate	S.E.	P-Value
Positive Climate	0.472	.039	0.684	.037	0.000
Negative Climate	-0.268	.030	-0.527	.048	0.000
Teacher Sensitivity	0.615	.033	0.940	.023	0.000
Regard for Student Perspectives	0.536	.036	0.784	.028	0.000

Note. All values significant at $p < .01$.

Table 15. STRQ CFA R-Squared Results

Observed Variable	R-Squared	S.E.	P-Value
Positive Climate	0.468	.051	0.000
Negative Climate	0.278	.051	0.000
Teacher Sensitivity	0.884	.042	0.000
Regard for Student Perspectives	0.615	.044	0.000

Note. All values significant at $p < .01$.

Primary Results

One model was used to test all three research questions. Absolute fit indices were not available for the primary model due to the latent interaction. The moderation of STRQ (i.e., latent factor) was tested between HS-STC variables and kindergarten reading outcomes. Thus, when a latent interaction was included in *Mplus* (Muthén & Muthén, 1998-2017) using the XWITH command, the absolute fit indices were no longer included in the output. See Appendix H for the primary analysis code.

Research Question 1: STRQ

Is STRQ formed in preschool significantly associated with kindergarten reading outcomes for students with disabilities and delays?

Hypothesis 1 predicted STRQ in preschool would be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between STRQ and kindergarten reading outcomes. The results indicated that STRQ in preschool was not significantly associated with kindergarten reading outcomes for students with disabilities and delays, with a standardized estimate $b = -0.204$ ($p = 0.161$) and an unstandardized estimate $B = -0.205$ ($p = 0.164$).

Research Question 2: Main Effects

Are HS-STC variables associated with kindergarten reading outcomes for students with disabilities and delays?

Hypothesis 2 predicted HS-STC variables would be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a statistically significant parameter estimate between HS-STC variables and student's kindergarten reading outcomes. After my dissertation proposal, I added Hypothesis 2a to the present study based on feedback from my dissertation committee that student gender is a salient variable related to student reading outcomes. The standardized results for each hypothesis are reported in text. See Table 16 for the unstandardized results.

Table 16. Unstandardized Main Effects between HS-STC Variables and Kindergarten Reading

Variable	Estimate	S.E	P-Value
Cohort	-0.108	0.144	0.455
Student Gender	**-0.396**	**0.120**	**0.001**
Student Race	0.052	0.222	0.816
Student SES	-0.161	0.121	0.184
Student Problem Behavior	0.089	0.077	0.248
HS Reading	**0.450**	**0.081**	**0.000**
Teacher Experience	**-0.157**	**0.061**	**0.010**
Teacher Mental Health	0.100	0.098	0.307
Class Size	0.002	0.063	0.971

Note. Bold = significant at $p < .05$.

Hypothesis 2a. *Student gender will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between student gender and student kindergarten reading outcomes.* The results indicated that student gender was significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate was b = -0.189 (p = 0.001), suggesting that being a male student was associated with poorer kindergarten reading outcomes.

Hypothesis 2b. *Student race will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between student race and student kindergarten reading outcomes.* The results indicated that student race was not significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate was b = -0.022 (p = 0.817), suggesting that student racial status was not significantly associated with kindergarten reading outcomes. Specifically, for students, a Black versus "Other" racial status (i.e., White, Hispanic/Latino, American Indian or Alaska Native, Asian or Pacific Islander, Multiracial/Biracial) was not significantly associated with kindergarten reading outcomes.

Hypothesis 2c. *Student SES will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between student SES and student kindergarten reading outcomes.* The results indicated that student SES was not significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The

standardized estimate was b = -0.080 (p = 0.177), suggesting that student SES was not significantly associated with kindergarten reading outcomes.

Hypothesis 2d. *Student behavior in preschool will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a negative statistically significant parameter estimate between student behavior ratings in preschool (i.e., lower behavior ratings indicative of positive behavior) and student kindergarten reading outcomes.* The results indicated that student behavior ratings by their teachers were not significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate was b = 0.088 (p = 0.240), suggesting that student behavior ratings in Head Start (i.e., lower behavior ratings indicative of positive behavior) were not significantly associated with kindergarten reading outcomes.

Hypothesis 2e. *Teacher experience in preschool will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a positive statistically significant parameter estimate between preschool teacher experience and student kindergarten reading outcomes.* The results indicated that teacher experience in Head Start was significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate was b = -0.156 (p = 0.009), suggesting that teaching experience in Head Start was significantly associated with kindergarten reading outcomes. Specifically, having a Head Start teacher with less experience was associated with higher kindergarten reading outcomes.

Hypothesis 2f. *Teacher mental health difficulties will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a negative statistically*

significant parameter estimate between preschool teacher mental health difficulties and student kindergarten reading outcomes. The results indicated that Head Start teacher mental health was not significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate was b = 0.078 (p = 0.303), suggesting that teacher mental health difficulties (i.e., teacher's self-report of depressive symptoms) was not significantly associated with kindergarten reading outcomes.

Hypothesis 2g. *Class size will be associated with kindergarten reading outcomes for students with disabilities and delays as indicated by a negative statistically significant parameter estimate between preschool class size and student kindergarten reading outcomes.* Finally, the results indicated that Head Start class size was not significantly associated with kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate was b = 0.002 (p = 0.971), suggesting that students' class size in Head Start was not significantly associated with kindergarten reading outcomes.

Research Question 3: Moderation

Does STRQ in preschool moderate the association between HS-STC variables and student's kindergarten reading outcomes for students with disabilities and delays?

Hypothesis 3 predicted that STRQ in preschool would moderate the association between HS-STC variables and kindergarten reading outcomes for students with disabilities and delays. Like Hypothesis 2a, I added Hypothesis 3a to the present study after the dissertation proposal based on feedback from my dissertation committee indicating that student gender is an important construct related to student reading outcomes. Hypothesis 3a examined the moderating role of

STRQ between student gender and kindergarten reading outcomes. The standardized results for each hypothesis are reported in text. See Table 17 for the unstandardized results.

Table 17. Unstandardized Interaction Effects between (STRQ/HS-STC Variables) and Kindergarten Reading

Variable	Estimate	S.E	P-Value
Gen*STRQ	**0.319**	**0.115**	**0.006**
Race*STRQ	-0.100	0.205	0.627
SES*STRQ	-0.086	0.139	0.538
ProBeh*STRQ	-0.099	0.086	0.252
TeachEx*STRQ	0.083	0.066	0.211
TeachMH*STRQ	0.079	0.091	0.385
ClassSize*STRQ	-0.039	0.050	0.435

Note. Gen = Gender; ProBeh = Student Problem Behavior; TeachEx = Teacher Experience; Teach MH = Teacher Mental Health. Bold = significant at $p < .05$.

Hypothesis 3a. *STRQ in preschool will moderate the association between student gender and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student gender and kindergarten reading outcomes when strong STRQ is present.* The results indicated that STRQ did significantly moderate the association between student gender and kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was $b = 0.152$ ($p = 0.004$), suggesting that STRQ did significantly moderate the association between student gender and kindergarten reading

outcomes. Specifically, when STRQ in Head Start improved, the association between student gender and kindergarten reading outcomes weakened. The moderating effect of STRQ served as a buffer such that male students tended to perform more favorably on kindergarten reading outcomes when STRQ in Head Start was high.

Due to the significant moderation effect, an interaction plot was created to visually examine the interaction between student gender and STRQ predicting kindergarten reading outcomes. See Appendix I for the interaction plot. The plot was created in Microsoft Excel using the STRQ summary score. The plot is organized with the outcome variable (i.e., kindergarten reading scores) on the y axis and STRQ on the x axis. The interaction is shown between male and female students with disabilities and delays, such that male students tend to fare better academically in kindergarten when STRQ is one standard deviation higher than the average STRQ in Head Start. Specifically, kindergarten academic scores for male students with disabilities and delays improved approximately eight and a half points when STRQ was higher in Head Start, which is more than half a standard deviation on the Woodcock Johnson Letter-Word Identification test. Thus, the plot illustrates that kindergarten reading outcomes are improved for male students with disabilities and delays when STRQ is higher in the preschool classroom.

The plot also illustrates that female students with disabilities and delays tend to have higher kindergarten outcomes compared to their male peers with disabilities when STRQ is considered average and low (i.e., one standard deviation below the norm in Head Start). Interestingly, for female students with disabilities and delays, kindergarten reading scores dropped approximately four points when STRQ in Head Start improved by one standard deviation above the average STRQ. For students with disabilities and delays, it is important to note that STRQ did not influence male or female students' kindergarten reading outcomes by

more than a standard deviation, which would be 15 points on the WJ Letter-Word Identification test.

Hypothesis 3b. *STRQ in preschool will moderate the association between student race and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student race and kindergarten reading outcomes when strong STRQ is present.* The results indicated that STRQ did not significantly moderate an association between student race and kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was b = -0.042 (p = 0.630), suggesting that STRQ did not significantly moderate the association between student race and kindergarten reading outcomes.

Hypothesis 3c. *STRQ in preschool will moderate the association between student SES and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student SES and kindergarten reading outcomes when strong STRQ is present.* The results indicated that STRQ did not significantly moderate an association between student SES and kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was b = -0.043 (p = 0.529), suggesting that STRQ did not significantly moderate the association between student SES and kindergarten reading outcomes.

Hypothesis 3d. *STRQ in preschool will moderate the association between student behavior in Head Start and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between student behavior ratings and kindergarten reading outcomes when strong STRQ is present.* The results indicated that STRQ did not significantly moderate an association between student behavior ratings and kindergarten

reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was b = -0.098 (p = 0.265), suggesting that STRQ did not significantly moderate the association between student behavior ratings in Head Start (i.e., lower behavior ratings indicative of positive behavior) and kindergarten reading outcomes.

Hypothesis 3e. *STRQ in preschool will moderate the association between teacher experience and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between teacher experience and kindergarten reading outcomes when strong STRQ is present.* The results indicated that STRQ did not significantly moderate an association between Head Start teacher experience and kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was b = 0.082 **(**p = 0.213), suggesting that STRQ did not significantly moderate the association between Head Start teacher experience and kindergarten reading outcomes.

Hypothesis 3f. *STRQ in preschool will moderate the association between teacher mental health difficulties and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between teacher mental health difficulties and kindergarten reading outcomes when strong STRQ is present.* The results indicated that STRQ did not significantly moderate an association between Head Start teacher mental health difficulties and kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was b = 0.061 (p = 0.365), suggesting that STRQ did not significantly moderate the association

between Head Start teacher mental health difficulties (i.e., teacher's self-report of depressive symptoms) and kindergarten reading outcomes.

Hypothesis 3g. *STRQ in preschool will moderate the association between class size and kindergarten reading outcomes for students with disabilities and delays as indicated by a significantly weaker association between class size and kindergarten reading outcomes when strong STRQ is present.* Finally**,** the results indicated that STRQ did not significantly moderate an association between class size and kindergarten reading outcomes for students with disabilities and delays, when controlling for student background variables. The standardized estimate of the latent interaction was b = -0.039 (p = 0.432), suggesting that STRQ did not significantly moderate the association between class size and kindergarten reading outcomes.

Supplementary Analyses

To better understand the moderating effect of STRQ for male students with disabilities and delays, supplementary analyses were conducted. Specifically, I examined the association between student gender and student behavior items. In the present study, student gender and teacher ratings of students' ability to concentrate and pay attention (r = 0.152; p = .004) and fidgeting (r = 0.138; p = .009) were both significantly correlated, suggesting that male students with disabilities and delays were rated by their teachers as having increased difficulty paying attention and fidgeting compared to their female peers with disabilities and delays. Student gender was also significantly correlated with teacher ratings of students' acting too young for their age (r = 0.160; p = .003), hitting or fighting with peers (r = 0.262; p < .001), and having temper tantrums (r = 0.146; p = .006), indicating that male students with disabilities were rated by teachers as having increased behavioral challenges in the classroom compared to their female peers with disabilities and delays. Finally, on average, male students with disabilities and delays

with strong STRQ (i.e., one standard deviation above the mean STRQ) were rated by teachers as having lower externalizing concerns (M = 8.48) compared to male students with disabilities and delays with lower STRQ (M = 9.11).

Results Summary

Collectively, the results indicated that STRQ significantly moderated the effect between student gender and kindergarten reading outcomes for students with disabilities and delays. Specifically, higher levels of STRQ during preschool tended to be more influential for male students related to their reading outcomes in kindergarten. Specifically, when the moderation effect was plotted, kindergarten reading scores for females with disabilities and delays decreased approximately four points when STRQ was high in Head Start, while kindergarten reading scores for male students with disabilities and delays increased by over eight points when STRQ was high in Head Start. Because the moderation effect is significant between STRQ and gender, the main effect between gender and kindergarten reading outcomes is conditional upon the moderator, STRQ. Thus, the significant conditional effect between student gender and kindergarten reading outcomes indicated that male status tended to be associated with poorer kindergarten reading outcomes for students with disabilities and delays, dependent upon STRQ in Head Start.

All other interaction effects in the primary model were not significant. This suggests that STRQ did not moderate the non-significant associations between HS-STC variables (i.e., student race, student SES, student behavior, teacher mental health, class size) and kindergarten reading outcomes for students with disabilities and delays. Additionally, STRQ did not moderate the significant effect between teacher experience and kindergarten reading outcomes. The significant main effect between teacher experience and kindergarten reading outcomes suggested that higher

kindergarten reading outcomes were associated with students who had Head Start teachers with less experience, not dependent on STRQ. Additionally, the control variable, Head Start reading outcomes, was significantly associated with kindergarten reading outcomes for students with disabilities and delays, such that higher Head Start reading outcomes were associated with higher kindergarten reading outcomes as indicated by a standardized estimate of $b = 0.451$ ($p < 0.001$). All other main effects were not significant. Finally, STRQ in preschool was not significantly associated with kindergarten reading outcomes for students with disabilities and delays.

CHAPTER 5: Discussion

Well-established evidence suggests that student-teacher relationship quality (STRQ) is associated with positive academic outcomes for typically developing students (Sabol & Pianta, 2012). However, less is known about the importance of STRQ for students with disabilities and delays, especially among preschool populations. Limitations are present among studies that have investigated the importance of STRQ for students with disabilities, such as using small sample sizes (Blacher et al., 2014), including specific disability categories (Eisenhower et al., 2007), and focusing on school-age or adolescent populations (Hamre & Pianta, 2001), which can make findings difficult to generalize to younger students with disabilities and delays. Many studies also overemphasize the connection between STRQ and student background characteristics, such as student race or student gender (Jerome et al., 2009), rather than investigating how STRQ may influence pre-existing associations between student, teacher, and classroom variables and student outcomes. This is problematic as it misses possible opportunities to streamline intervention efforts via improving STRQ in the classroom. Finally, STRQ is commonly measured using an informant rating scale completed by teachers (e.g., Koenen et al., 2021), which limits the understanding of STRQ to only the teacher's perception of STRQ, which may be a somewhat biased representation of the construct.

The present study aimed to address these gaps by investigating the moderating role of STRQ between Head Start Student, Teacher, and Classroom (HS-STC) variables and kindergarten reading outcomes for preschool students with disabilities and delays. Specifically, my dissertation study used secondary data from the Family and Child Experiences Survey (FACES) 2009 Cohort, which included an observational measure of STRQ, the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) Emotional Support Domain. The

primary analysis included one latent-interaction structural equation model completed using *Mplus* Version 8.6 (Muthén & Muthén, 1998-2021) software. The primary results suggested that STRQ moderated the association between student gender and kindergarten reading outcomes for students with disabilities and delays. Additional findings indicated that student gender, teacher experience, and Head Start reading scores were significantly associated with kindergarten reading outcomes for students with disabilities and delays.

The primary research questions were tested using one model, which was advantageous. I interpreted my results while accounting for all HS-STC variables, as the main effects and latent interactions were tested simultaneously while including all variables of interest. If I ran multiple models, such as one model with student variables, one model with teacher variables, and one model with the classroom variable, I would have been unable to understand the role of STRQ when all variables are accounted for simultaneously. Rather, the results for student variables, teacher variables, and the classroom variable would have been interpreted in isolation, which does not align with how they naturally occur in the school setting, or with EST. Using more than one model would have also increased the number of statistical tests and chance of error in the model computation. I also avoided using model building, such as running one model for each research question, in an effort to understand the importance of STRQ for students with disabilities and delays when all variables were considered at once.

Overall, my dissertation study adds to the current literature as it included an observational measure of STRQ (i.e., CLASS; Pianta et al., 2008), investigated the importance of STRQ for preschool students with disabilities and delays, and examined student, teacher, and environmental factors related to student academic outcomes. My study findings provided some evidence that STRQ in preschool could buffer the association between student gender and future

academic outcomes for students with disabilities and delays, which can inform future research and intervention efforts for young students with disabilities and delays.

Summary of Measurement Model

Prior to running the primary analysis, a confirmatory factor analysis (CFA) was completed for the variable of interest, STRQ, using the CLASS Emotional Support Domain. Adequate to moderate model fit resulted. All four CLASS indicators (i.e., Positive Climate, Negative Climate, Teacher Sensitivity, Regard for Student Perspectives) significantly loaded onto the STRQ factor, which aligns with Pianta's (2008) proposed CLASS structure and with current research that has examined the CLASS factor structure (Li et al., 2020). Therefore, my final model included STRQ as a latent factor with four indicators as suggested by both research and theory.

After completing a series of CFAs to examine the factor structure for the student behavior variable (i.e., teacher report of student behavior), a summary score including only externalizing behaviors was included in the model. The externalizing behaviors included: 1. Age-appropriate behavior; 2. Difficulty concentrating and maintaining attention; 3. Hitting or fighting; 4. Fidgeting; difficulty sitting still; 5. Temper tantrums (USDHHS, 2009a). This differed from the FACES 2009 study that included all behavioral items (USDHHS, 2009a), which should be considered when comparing my study results to other research that uses FACES 2009 data. My decision to include only student externalizing behaviors was based on the CFA model fit and factor loadings, which suggested that the student externalizing items loaded strongly onto the factor with good model fit. This finding may be evidence of student externalizing behaviors being easier for teachers to identify, compared to internalizing behaviors. For example, in a study of preschool students, teachers' ratings of externalizing behaviors were predictive of observed

behaviors, but this was not true for their ratings of student internalizing concerns (Hinshaw et al., 1992).

STRQ Moderation and Main Effects

Overview of Findings

The primary results answered my three research study questions, which examined 1. Is STRQ formed in preschool significantly associated with kindergarten reading outcomes for students with disabilities and delays?; 2. Are HS-STC variables associated with kindergarten reading outcomes for students with disabilities and delays?; and 3. Does STRQ in preschool moderate the association between HS-STC variables and student's kindergarten reading outcomes for students with disabilities and delays? All research questions were tested using one statistical model. STRQ, the variable of interest, was measured using the CLASS Emotional Support Domain. I made the decision to use one CLASS domain based on multiple studies indicating a strong correlation between the CLASS Emotional Support Domain and the Student Teacher Relationship Scale (Pianta & Steinberg, 1992), which is a commonly used measure of STRQ (Moen et al., 2019; Walker & Graham, 2021). In addition, the CLASS Emotional Support Domain aligns with the definition of STRQ in the literature (Rose et al., 2019).

Research question one examined if STRQ was associated with kindergarten reading outcomes for students with disabilities and delays. I predicted that STRQ in Head Start would be associated with kindergarten reading outcomes for students with disabilities and delays. The results indicated that STRQ in Head Start was not significantly associated with kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) for students with disabilities and delays.

Research question two examined if HS-STC variables were associated with kindergarten reading outcomes for students with disabilities and delays. I predicted that HS-STC variables (i.e., student gender, student race, student SES, student behavior, teacher experience, teacher mental health, and class size) would be associated with kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) for students with disabilities and delays, while controlling for student background variables (i.e., student cohort, student Head Start reading outcomes). Three significant correlations resulted in the primary analysis, specifically between 1. Student gender and kindergarten reading outcomes, 2. Teacher experience and kindergarten reading outcomes, and 3. Head Start reading outcomes and kindergarten reading outcomes.

Research question three examined if STRQ in preschool moderated the association between HS-STC variables and student's kindergarten reading outcomes for students with disabilities and delays. I predicted that STRQ would moderate the association between HS-STC variables and kindergarten reading outcomes for students with disabilities and delays, while controlling for student background variables. Regardless of a significant main effect between HS-STC variables and kindergarten reading outcomes, moderation was considered for all associations because HS-STC variables and kindergarten reading outcomes may not have been significant when STRQ was at average levels of STRQ for the sample, but the association may have changed if STRQ increased or decreased. Hence, the association between HS-STC variables and kindergarten reading outcomes may have been conditional on STRQ for students with disabilities and delays.

The moderating effect of STRQ was tested in the primary model by way of a latent interaction between STRQ and each HS-STC variable while controlling for student background variables. Unlike many other studies that have used observed variables to measure STRQ (e.g.,

Blacher et al., 2014; Fowler et al., 2008), a latent term was used in the primary model to increase model precision regarding the variable of interest, STRQ. The present study findings suggested that STRQ significantly moderated the effect between student gender and kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) for students with disabilities and delays. All other moderation effects were not significant. The findings from all research questions are discussed integrally below.

Student Gender

STRQ moderated the effect between student gender and kindergarten reading outcomes for students with disabilities and delays. When STRQ in Head Start improved, the association between student gender and kindergarten reading outcomes weakened for students with disabilities and delays. Specifically, the moderating effect of STRQ served as a buffer such that male status was less strongly associated with poorer kindergarten reading outcomes when STRQ in preschool was strong. In other words, male students with disabilities and delays tended to have poorer letter-word identification outcomes compared to female students with disabilities and delays. However, STRQ moderated this association as kindergarten letter-word identification outcomes improved for male students with disabilities at higher levels of STRQ, while female students with disabilities performed poorer on kindergarten letter-word identification outcomes at higher levels of STRQ.

Generally, kindergarten reading scores did not vary much for female students with disabilities compared to males with disabilities, which may be suggestive as to why high STRQ may not have been as meaningful for this group. Specifically, kindergarten letter-word identification scores dropped approximately four points for female students with disabilities who were in Head Start classrooms with higher STRQ, which are similar letter-word identification

scores when STRQ in Head Start classrooms was low. Thus, the present study findings should be considered such that STRQ changed very little for the female students with disabilities.

Other factors, besides STRQ, may be more influential for female reading outcomes. For instance, Nalipay and colleagues' (2020) study provides evidence that parental emotion about reading can influence their child's emotion about reading, subsequently affecting their reading outcomes. Thus, parental attitudes may also play a role in reading outcomes for female students. An alternative rationale could be that at higher levels of STRQ, female students may have been more distracted by student-teacher interactions, which may have contributed to the decreased reading outcomes for female students with disabilities and delays. Thus, a variety of student, familial, and/or classroom factors may have influenced the present study outcomes for female students with disabilities and delays.

Male students who tend to perform lower on reading outcomes during early schooling may especially benefit from positive STRQ compared to female students who tend to perform higher on reading outcomes (Logan & Johnston, 2010). Students who experience most academic risk are likely to benefit the most from positive STRQ. In the United States, a reading gender gap has been identified among school-aged students, in which female students tend to perform higher than their male peers (Reardon et al., 2019). A recent study of early elementary students found that STRQ predicted reading outcomes for boys, while STRQ predicted math outcomes for girls (Valiente et al., 2019). Thus, findings may have differed if another academic subject were examined. Overall, it is important to note that kindergarten reading scores for female students with disabilities and delays did not fluctuate much in the present study despite differences in Head Start STRQ.

Student gender was significantly associated with kindergarten reading outcomes for students with disabilities and delays, contingent on STRQ. Specifically, identifying as a male student with a disability or delay was associated with poorer kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) as female students with disabilities and delays tended to outperform male students with disabilities and delays on kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) at average levels of STRQ for the sample. Due to the significant moderating effect, the association between student gender and kindergarten reading outcomes differed depending on STRQ in Head Start. Namely, female students with disabilities and delays tended to have better kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) when STRQ was average. However, when STRQ in preschool was strong, male students with disabilities tended to have better kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) compared to their female peers with disabilities and delays. Thus, the significant main effect between student gender and kindergarten reading outcomes was contingent upon the level of STRQ for students with disabilities and delays. The significant main effect is supported by the vast literature that suggests that female students tend to outperform male students on reading achievement tests (e.g., Liederman et al., 2005; Logan & Johnston, 2010).

Positive STRQ may be particularly important for male students with disabilities as this group oftentimes experiences poorer STRQ. Multiple studies suggest that male students (Jerome et al., 2009; McGrath & Bergen, 2015; Rudasill et al., 2010) and students with disabilities (Blacher et al., 2009; Freire et al., 2020; McGrath & Bergen, 2015; Prino et al., 2016) experience poorer STRQ compared to their female, typically developing peers. While male students with disabilities seem particularly vulnerable to experiencing low STRQ, it may in fact be a

particularly important protective factor in the promotion of academic outcomes. The significant moderating effect of STRQ in the present study suggests that male students with disabilities and delays performed better on kindergarten academic outcomes at higher levels of positive STRQ in preschool. Thus, for male students who are at an increased risk of experiencing negative STRQ and poorer reading outcomes, the effects of high STRQ may be especially meaningful for future academic outcomes.

Much of the research on the importance of STRQ was completed with typically developing students and their teachers in the early and late 2000s, and Robert Pianta and Bridget Hamre were well known for their work in this area (e.g., Hamre & Pianta, 2001; Hamre & Pianta, 2004; Pianta et al., 2012; Pianta & Stuhlman, 2004). Current research efforts investigating STRQ need to be extended to young students with disabilities. Despite few studies that have investigated the intricacies of STRQ for male students with disabilities, some studies have explored the importance of STRQ for young students and at-risk populations. McGrath and Van Bergen (2015) emphasize the importance of STRQ during early childhood as STRQ during a student's early education has been associated with later academic success, engagement, and adjustment among students. Research also suggests that STRQ is positively associated with academic outcomes for young students who present with multiple risk factors (Cadima et al., 2010). Cadima and colleague's (2010) findings are relevant in the present study as students with disabilities (i.e., at-risk students) and male students tend to experience lower academic performance compared to their typically developing peers (Gilmour et al., 2019; Logan & Johnston, 2010). Thus, for young students with multiple risk factors, such as male students with disabilities, strong STRQ can serve as a protective factor related to future academic outcomes.

The significant moderating effect of STRQ for male students with disabilities is particularly salient when considering contextual risk factors, as my sample of Head Start students tended to have lower household incomes compared to the US national average. For instance, the median household income for my Head Start sample was between $15,000 - $20,000, while the U.S. national average in 2009 was a little over $51,000 (USDHHS, 2019; U.S. Census Bureau, 2011). Less than 10% of my sample lived in families with incomes higher than the national average. Students living in low-income families are at greater risk for additional risk factors, such as inconsistent housing, living in unsafe neighborhoods, higher school drop-out rates, less access to adequate nutrition, and increased developmental, behavioral, and physical health needs (American Psychological Association, 2022). In a recent study by Hynek and colleagues (2022), preschool children living in low-income settings for an extended period were at an increased risk of needing outpatient mental health services when they are older. In addition to these risk factors that can each negatively impact academic outcomes, students living in poverty tend to have poorer academic outcomes compared to their higher income peers (Hanushek et al., 2019).

Contextual factors related to my sample, such as low-income status, are meaningful as it informs the implications of the findings. Compared to their more privileged peers, male students with disabilities may especially benefit from positive relationships with their teachers as positive STRQ may serve as a malleable protective factor amidst other compounding risk factors, such as having a disability and living in a low-income household. Thus, cultivating strong STRQ for male students with disabilities who experience various academic and contextual risk factors during early childhood may be especially salient for their future academic outcomes.

STRQ may have been particularly important for male students with disabilities due to confounding factors that can adversely impact educational outcomes for male students compared

to their female peers. Specifically, STRQ may have moderated the association between student gender and kindergarten reading outcomes based on research that suggests that positive STRQ can support students' attention difficulties, externalizing behaviors, and language development. This is meaningful as confounding factors such as these may all adversely contribute to differences in reading outcomes for male students with disabilities (Adani & Cepanec, 2019; Paz et al., 2020). Thus, STRQ may have boosted academic outcomes for male students with disabilities by serving as a protective factor amid areas in which male students with disabilities tend to have heightened needs (i.e., attention, externalizing behaviors, language development), which can all adversely impact academic performance.

First, according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), two male students are diagnosed with attention deficit hyperactivity disorder (ADHD) for everyone one female (American Psychiatric Association, 2013). Students with attention challenges can experience increased rates of reading challenges, as ADHD and reading disorders (i.e., dyslexia) can be comorbid disorders (Shaywitz & Shaywitz, 2008). In a study of twins with and without reading disabilities, students with reading disabilities were more likely to be diagnosed with ADHD (Willcutt & Pennington, 2000). In the present study, both major ADHD diagnostic criteria, inattention and hyperactivity, were significantly associated with reading disabilities for male students. Male students with disabilities and delays were rated by their teachers as having increased difficulty paying attention and fidgeting compared to their female peers with disabilities and delays, which aligns with recent research.

Research suggests that STRQ is meaningful in supporting students' attention. In a study of 140 preschool students, males were perceived as having poorer STRQ than female peers by their teachers (Acar et al., 2022). Negative STRQ was negatively associated with students'

learning behaviors, which included students' attention, motivation, and attitudes about learning. This finding was meaningful as learning behaviors decreased by over six points (scale of 14-58) when negative STRQ increased by one-point (five-point scale). This applies to the current study as positive STRQ in the preschool classroom may have played a role in fostering positive learning behaviors among male students with disabilities and delays, which could have contributed to the moderating effect on kindergarten reading outcomes.

Similar to attention concerns, externalizing behaviors have been found to negatively influence academic performance among students with disabilities (Reid et al., 2004). Male students tend to exhibit increased rates of externalizing behaviors, recognized as early as toddlerhood (Paz et al., 2020). In my dissertation study, student gender was significantly correlated with teacher ratings of students' acting too young for their age, hitting or fighting with peers, and having temper tantrums. The findings suggest that teachers rated male students with disabilities and delays higher in externalizing behavior challenges compared to their female peers with disabilities and delays, which may have contributed to poorer kindergarten reading outcomes, as externalizing behavioral challenges are often associated with poorer academic outcomes (Hinshaw, 1992; Kulkarni et al., 2020; Okano et al., 2020).

STRQ may have moderated the association between student gender and kindergarten reading outcomes due to the importance of positive STRQ for minimizing externalizing behaviors among male students with disabilities and delays. Using data from over 1,000 students who participated in the NICHD Study of Early Child Care and Youth Development study, O'Connor and colleagues (2011) found that positive STRQ predicted fewer externalizing concerns among students. Likewise, negative STRQ was a stronger predictor of externalizing concerns for male students compared to their female peers in a study of over 300 Head Start

students (Ewing & Taylor, 2009). Even more, in my dissertation study, on average, male students with disabilities and delays who had strong STRQ were rated by teachers as having lower externalizing concerns compared to male students with poorer STRQ. Thus, positive STRQ may have supported favorable academic outcomes for male students with disabilities and delays by also buffering extraneous factors, such as student externalizing problems.

Finally, nearly 70% of my sample were identified as having a speech impairment. Current research suggests that male students are found to have higher rates of speech and language disorders compared to female peers (Adani & Cepanec, 2019). Positive STRQ has been connected to improved language skills for at-risk preschool students (Schmitt et al., 2012), which is meaningful as proficient language skills are associated with higher academic performance among young children (Kastner et al., 2001). Thus, for male students with a speech impairment or delay, STRQ may have inadvertently supported kindergarten academic outcomes via language development in preschool. Overall, STRQ may have bolstered academic outcomes for male students with disabilities by serving as a protective factor amongst areas in which male students with disabilities tend to have increased needs (i.e., attention, externalizing behaviors, language development), which can all adversely impact educational performance.

Although most students in my sample were identified as having a speech and language impairment, many students may have experienced comorbid disabilities or delays that were not recognized. For instance, studies have found that speech impairments commonly overlap with ADHD diagnoses (Tomblin & Mueller, 2012). Yet, children living in low-income households, like my sample, can experience barriers to accessing mental health care (American Academy of Pediatrics, 2016) which can pose challenges to receiving a diagnosis and appropriate care. For instance, developmental concerns related to autism spectrum disorder are often present within

the first few years of life (CDC, 2019a), yet socioeconomic disparities tend to delay the diagnosis of autism spectrum disorder for children living in poverty (Durkin et al., 2010), which can prolong access to early intervention services. Thus, compared to their more privileged peers, at-risk populations can experience limited resources and may not receive evidence-based treatments (Castro-Ramirez, 2021). Along with living in low-income settings, my study sample were newly enrolled Head Start students aged three to five years old. Because the median age of onset for most mental health diagnoses ranges from school-age to adulthood (Solmi et al., 2022), many mental health needs may not have been recognized in my early childhood population. In addition, it is possible that teachers may not have known if a student had a disability in the fall of 2009 as students were newly enrolled into the program (USDHHS, 2019). Thus, more nuanced needs, delays, and disabilities may have been present within my sample that were not identified due to contextual and demographic sample characteristics. Therefore, when interpreting the importance of STRQ for male students, additional needs may have been present that were not captured in the data.

The significant moderating effect of STRQ between student gender and kindergarten reading outcomes has strong implications for interventions, particularly for male students with disabilities who tend to fare better academically when STRQ in preschool is strong. Because STRQ is a malleable factor, improving STRQ can be a targeted point of intervention that buffers the impact of fixed student characteristics, such as gender, on future academic outcomes for students with disabilities. Improving STRQ in the classroom could serve as a protective factor and indirectly support academic outcomes for male students with disabilities. Thus, the significant moderation that resulted between student gender and STRQ is especially meaningful as targeted intervention strategies to improve STRQ are possible in the school context.

Increased efforts to improve STRQ are needed based on current trends in the literature that report both male students and students with disabilities tend to experience poorer STRQ compared to their female and typically developing peers. For instance, Prino and colleagues (2016) found that students with disabilities, such as students with autism or ADHD, tend to experience poorer STRQ compared to their typically developing peers. Additionally, multiple studies report male students tend to experience poorer STRQ than female students (Jerome et al., 2009; McGrath & Bergen, 2015; Rudasill et al., 2010). Thus, intentional efforts to improve STRQ, particularly for male students with disabilities, are needed. Current school practices, such as Positive Behavioral Interventions and Supports (PBIS; Center on PBIS, 2022), provide a framework that emphasizes the importance of establishing positive relationships between students and teachers universally. However, in addition to Tier 1 interventions, targeted interventions are needed to address the gap in STRQ for male students with disabilities and delays as high quality STRQ could have secondary benefits for students' future academic outcomes.

Strategies aimed at supporting student-teacher relationships for preschool students are common practice. Yet, like practices for school-aged students, evidence-based interventions aimed at supporting STRQ specifically for preschool students with disabilities are limited. For example, the Center on the Social and Emotional Foundations for Early Learning (CSEFEL) provides teachers with training modules and resources, such as examples about how to give positive feedback to students, in efforts to support preschool teachers in cultivating positive STRQ with their students (CSEFEL, 2022). The literature highlights teaching strategies that aim to boost student-teacher relationships, such as providing students with praise and taking time to get know each student (Rimm-Kauman & Sandilos, 2015). However, researchers have identified

the need for more specific intervention efforts for teachers of students with disabilities, such as teacher-student interaction coaching (Koenen et al., 2021).

Although evidence-based interventions specific to teachers and students with disabilities are lacking, Teacher-Child Interaction Training – Universal (TCIT-U; Gershenson et al., 2020) is one approach that has been used to improve student-teacher relationships in the classroom. TCIT-U was created using similar principles as Parent-Child Interaction Therapy (PCIT; Eyberg, 1988), an evidenced-based intervention for caregivers and children with challenging behavior. TCIT-U focuses on enhancing student-teacher relationships in the classroom by providing professional development to educators. Research suggests that TCIT-U can be useful for all students, including those with disabilities and delays, as it teaches educators behavioral strategies and relationship skills to use with their students (Stern & Budd, 2021).

Although evidence-based interventions aimed at supporting positive STRQ specifically for male preschool students with disabilities are limited, Murray and Pianta (2007) highlight four interdependent factors that contribute to STRQ and can serve as a framework for designing future intervention. The four factors include: 1. Organizational Structures and Resources; 2. Classroom Structures and Practices; 3. Teacher Beliefs, Behaviors, and Actions, and 4. Individual Skills for Developing Prosocial Relationships. First, *Organization Structure and Resources* to promote positive STRQ can include inclusionary practices and extended time spent with the same teacher. Second, *Classroom Structures and Practices* can include explicit teaching of the classroom rules and expectations to create a safe, predictable, and trusting environment for students to learn. Third, *Teacher Beliefs, Behaviors, and Actions* can include holding students to high expectations that are appropriate to their developmental level, teachers examining their own beliefs and how this influences their teaching practices, teachers providing positive praise and

feedback to students, and teachers taking time to get to know students. Finally, *Individual Skills for Developing Prosocial Relationships* can include teaching students components of social-emotional learning to help them develop interpersonal skills needed to form strong relationships with adults and other students (Murray & Pianta, 2007). Although this framework was designed for adolescents with disabilities, it emphasizes the multi-faceted approach that researchers are encouraged to consider when developing interventions to improve STRQ for young students with disabilities. It also provides a guide for schools and practitioners to generalize for younger students with disabilities.

Established interventions aimed at supporting STRQ address the concepts outlined by Murray and Pianta (2007). For instance, The Incredible Years Program (Webster-Stratton, 2005) is an evidence-based program that aims to enhance school-related outcomes for students while supporting high quality relationships between students and their parents and teachers. The Incredible Years Program includes Child, Parent, and Teacher programs, which all aim to foster students' social-emotional skills, which encompasses Murray and Pianta's (2007) *Individual Skills for Developing Prosocial Relationships*. Even more, the Incredible Beginnings Program, an adapted version of the Incredible Years program for teachers of younger students, teaches developmentally appropriate teaching strategies rooted in developing positive STRQ, which aligns with Murray and Pianta's (2007) *Teacher Beliefs, Behaviors, and Actions*. Thus, established interventions, like the Incredible Years Program, incorporate strategies to enhance STRQ and have also been adapted for teachers and parents working with children with autism spectrum disorder. Additional adaptations of existing programs, such as TCIT and the Incredible Years Program, are needed particularly for male students with disabilities.

Teacher Experience

Next, the present study findings indicated that teacher experience was associated with kindergarten reading outcomes for students with disabilities and delays, while controlling for student background variables. Specifically, having a Head Start teacher with less experience was associated with higher kindergarten letter-word identification outcomes (i.e., kindergarten reading outcomes) for students with disabilities and delays. Little research was found related to teacher experience and student outcomes for preschool students with disabilities and delays.

Inconsistent findings emerge in the literature on typically developing students. For instance, when using the Early Childhood Longitudinal Data Set (ECLS-K), Croninger and colleagues (2007) suggest that students who had teachers with less than two years of experience tended to have poorer reading outcomes. In a more recent review article of 30 studies, Podolsky and colleagues (2019) reported that more experienced teachers tended to have students who perform better academically. Conversely, in a study by Connor and colleagues (2005), less experienced teachers tended to have students who performed better in early reading. Thus, research findings related to teacher experience and student academic outcomes differ.

My dissertation study contributes to the lack of research in the disability literature related to the importance of teacher experience for students' reading outcomes. First, my sample of preschool students with disabilities differs from published research studies conducted with school-aged typically developing students. The importance of teacher experience may differ for students of different ages and ability statuses. Yet, upon a thorough literature search, there were no studies found that have investigated the differential importance of teacher experience for students with and without disabilities across multiple ages, which provides ample opportunity for future research.

Second, my dissertation study used a Head Start sample, which may have influenced the finding that students with teachers with less experience tended to have higher kindergarten outcomes, as Head Start teachers utilized Head Start curricula, which includes an evidence-based curriculum, professional development opportunities, and teacher mentoring (USDHHS, 2021; USDHHS, 2022) - all of which may support teachers with less experience. Researchers are encouraged to investigate if the importance of teacher experience on future academic outcomes for preschool students with disabilities differs based on the type of early childhood program, such as federally funded preschools (i.e., Head Start), private preschool programs, or in-home preschools. This is important as confounding factors associated with the different programming, such as teacher professional development, teacher support, and evidence-based curricula, could influence how salient teacher experience may be for positive academic outcomes for students with disabilities and delays. For example, Head Start teachers who engaged in professional development consisting of mentoring, self-reflection, workshops, and peer-to-peer coaching were found to have significantly better teaching practices (e.g., productivity, behavior management, student feedback, etc.) compared to teachers who did not participate (Zan & Donegan-Ritter, 2014). These findings were similar for teachers who reported having and not having a college degree.

In my study, Head Start teachers all had teacher assistants in their classrooms. Martin and colleagues (2021) suggest that teaching assistants who work in classrooms with students with disabilities that are more adaptable experience increased enjoyment, motivation, and participation in their work, which can all have meaningful effects for students. Thus, characteristics of teaching assistants in the classroom may have lessened the need for primary teachers to have a wealth of experience. Future studies are encouraged to examine how the

characteristics of all classroom teachers may influence outcomes for students with disabilities and delays.

Finally, research suggests that teachers of students with disabilities with more experience tend to experience higher rates of teacher burnout (Williams & Dikes, 2015). Higher rates of stress and burnout are associated with poorer student outcomes (Herman et al., 2018). If Head Start teachers with more experience in my sample had higher rates of stress and burnout, this could have negatively impacted their teaching, which in turn could influence student learning and academic outcomes. When investigating the importance of teacher experience for students with disabilities and delays in future studies, researchers are encouraged to include measures of teacher stress and burnout, as these factors are particularly salient for teachers of students with disabilities (Brunsting et al., 2014), and ultimately could be related to academic outcomes for students with disabilities and delays. For instance, teachers who experience burnout and stress may be less warm and responsive to students, which were identified as important characteristics for students' future reading outcomes in a study by Connor and colleagues (2005).

Head Start Reading

Finally, one student background variable, student Head Start letter-word identification was positively associated with student kindergarten letter-word identification outcomes for students with disabilities and delays. This aligns with the literature for typically developing populations that suggests that current academic performance is predictive of later academic outcomes (Croninger et al., 2007; Duncan et al., 2007). These findings are also consistent for preschool students as academic growth in Head Start has been found to be associated with higher kindergarten outcomes (Ramsook et al., 2020). Therefore, the results from this dissertation study

suggests that trends for students with disabilities and delays align with that of their typically developing peers such that earlier academic outcomes are predictive of later academic outcomes.

Other HS-STC Variables

Unlike the significant moderating effect of STRQ on student gender and reading outcomes, STRQ did not moderate an association between any other HS-STC variables and students' kindergarten reading outcomes. These findings differ from studies that suggest STRQ may be more closely related to academic outcomes compared to HS-STC variables (e.g., Zendarski et al., 2020). For instance, in a study of students with and without attention deficit hyperactivity disorder (ADHD), Zendarski and colleagues (2020) found that students' reading scores were more closely related to STRQ compared to student variables, such as student SES. Similarly, student reading scores were also more closely associated with STRQ compared to teacher variables, such as teacher experience.

The findings from this dissertation study may contrast results in the current literature due to the STRQ measure used in my study. Because I used a class-level measure of STRQ in my dissertation, I was unable to capture the importance of each student's individual relationship quality formed with their teacher. Rather, my study alluded to the importance of the STRQ between the teacher and all students in the classrooms. This measurement approach poses challenges as students with disabilities may have different STRQ than their typically developing peers (Blacher et al., 2009; McGrath & Bergen, 2015). While there were no level two effects in my study, suggesting that the global STRQ estimate was a good representation of the STRQ for all students in the classroom, future studies should investigate the nuances of STRQ for students with disabilities using a multi-dimensional approach. A comprehensive measurement approach using both an observational and informant-report measure of STRQ is recommended. This

approach could provide additional information about how class-level STRQ and each student's individual relationship with their teacher could buffer or enhance preexisting associations between HS-STC variables and kindergarten outcomes for students with disabilities and delays. Hence, a non-significant moderating effect may have resulted between many HS-STC variables and kindergarten reading outcomes based on the global STRQ measure in this study.

The CLASS Emotional Support Domain was selected as a measure of STRQ as it tends to be most strongly correlated with STRQ informant-reports, such as the Student Teacher Relationship Scale (Moen et al., 2019; Walker & Graham, 2021). Thus, selecting the CLASS Emotional Support Domain as a proxy for STRQ was beneficial as the four Emotional Support dimensions (i.e., Positive Climate, Negative Climate, Regard for Student Perspectives, Teacher Sensitivity) aligned with the current STRQ definition posed in the literature (Rose et al., 2019). However, including only one of the three CLASS domains may be another reason why STRQ was not found to moderate the association between HS-STC variables and kindergarten reading outcomes. By not including all CLASS domains, I may have missed important teacher practices implemented in the classroom that could help facilitate positive STRQ for students with disabilities. For example, teacher attributes measured by other CLASS domains, such as student feedback, behavior management skills, and language modeling may influence STRQ. Thus, the unidimensional approach to STRQ measurement in my study may have affected the moderation outcomes.

Additionally, STRQ may not have moderated the association between other HS-STC variables and kindergarten reading outcomes due to the use of a Head Start sample. Head Start emphasizes the importance of student-teacher relationships by embedding relationship-building practices in the programming (USDHHS, 2019). Thus, STRQ was relatively high across

classrooms, and there was little variability across classrooms. The lack of variability may have made it challenging to recognize differences between classrooms with low and high STRQ in relation to HS-STC variables and kindergarten reading outcomes for students with disabilities and delays, as very few classrooms would be characterized as having negative STRQ. Future researchers are encouraged to investigate STRQ's moderating role between student, teacher, and classroom variables and future academic outcomes among preschool populations with disabilities when differences in STRQ are more evident.

Additionally, all other main effects between HS-STC variables and kindergarten reading scores were not significant. Particularly, student race, student SES, student behavior, Head Start teacher mental health difficulties, and Head Start class size were not significantly associated with kindergarten reading outcomes for students with disabilities and delays. These findings differ from other studies that indicate student minority racial status (Wei et al., 2011), students with low SES (Hanushek et al., 2019), increased student problem behavior (Kulkarni et al., 2020), increased teacher-reported mental health concerns (McLean & Conner, 2015) and larger class sizes (Francis & Barnett, 2019) are associated with poorer academic outcomes.

My study findings may differ from previous research studies due to my use of a sample of preschool students with disabilities and delays. Typically developing students tend to perform better academically compared to students with disabilities. In a meta-analysis of 23 studies, students with disabilities were found to perform over one standard deviation lower on reading achievement measures compared to their typically developing peers (Gilmour et al., 2019). Thus, additional factors, such as student minority status, student SES, student problem behavior, teacher mental health, and class size may not be as meaningful for students with disabilities who may already be underperforming academically.

Most students in my sample had a disability or delay characterized as a speech impairment. It has long been recognized that preschool students with disabilities and delays benefit when they are integrated in inclusive settings (Odom, 2000). Head Start classrooms can provide a language-rich environment that supports student development, especially for students with disabilities. Wasik and colleagues (2006) report positive effects on student language when Head Start teachers incorporate language activities into academic tasks, such as providing feedback, modeling rich language, asking questions, and teaching listening skills. Additionally, young students with disabilities who are surrounded by typically developing peers with strong language skills can benefit. In a study of 670 preschoolers with and without disabilities, students' language skills in the spring were predicted by classmates' language skills in the fall (Justice et al., 2014). Classmate language skills were the most influential for the language development among students with disabilities, such that students with disabilities made the most gains when surrounded by peers with strong language skills (Justice et al., 2014). This is meaningful for preschool populations compared to school-aged students, as preschool students generally have more time to interact with one another, which can foster language skills through peer modeling, play, and practice. Thus, unlike studies that suggest smaller class sizes may be beneficial for academic outcomes (e.g., Francis & Barnett, 2019), the opposite may be true for young students with disabilities and delays who can learn additional skills from their peers during structured and unstructured activities.

Other characteristics of my study sample may have also contributed to the insignificant findings. For example, students in my study may live in households with lower SES compared to other preschool samples in the United States due to Head Start admission criteria that suggests students must be at or below the poverty line to receive services (USDHHS, 2022). Exceptions to

this criterion include students who are homeless, in foster care, or in a family receiving certain public assistance. Thus, although there was some variability in student SES, this may not mirror other preschool populations with no poverty criteria. For instance, over 80% of my Head Start sample had an average household income that was less than $35,000, while the U.S. national average was approximately $51,000 in 2009 (USDHHS, 2019; U.S. Census Bureau, 2011). Thus, the impact of SES on kindergarten reading outcomes for students with disabilities and delays may be more noticeable when there are more extreme differences in SES between students, such as for preschool students with disabilities who do not attend a federally funded early childhood program like Head Start.

Another consideration related to the difference in findings could be attributed to the variables themselves. For instance, most teachers in my sample did not report significant mental health concerns. Specifically, 61% percent of teachers reported not being depressed. Reports of student problem behavior were also relatively low across the sample, suggesting that many students were perceived by their teachers as using age-appropriate behaviors in the classroom. In addition, student SES was measured dichotomously (i.e., at/above poverty threshold, below poverty threshold), which may not have captured variability between different SES groups. Similarly, my race variable was measured dichotomously (i.e., Black, Other Races) due to the complexity of the primary model. Overall, different findings may have emerged if there were more variability in the sample variables.

Therefore, in future studies, researchers are encouraged to examine when HS-STC variables become salient for reading outcomes for preschool students with disabilities by including diverse samples and data with more variability. Examining at what degree teachers' report of mental health difficulties, student behavior ratings, or student SES significantly

influence future academic outcomes for students with disabilities and delays would propel the literature forward. One way to accomplish this aim is to use variables that capture the nuances of the sample, such as variables with increased variability, like a non-dichotomous variable for race or continuous variables for teacher mental health and SES. Additionally, including a preschool sample with more diverse student qualities (e.g., disabilities, SES, racial statuses, behavioral needs), teacher qualities (e.g., mental health needs, stress, experience, etc.), and classroom characteristics (e.g., size, resources, curriculum) could help clarify when HS-STC variables become especially important for students with disabilities' future academic outcomes. Finally, researchers should add a comparative sample of typically developing peers to understand how HS-STC variables may relate differently to academic outcomes for preschool students with and without disabilities and delays.

STRQ & Kindergarten Reading Outcomes

STRQ in Head Start was not significantly associated with kindergarten reading outcomes for students with disabilities and delays. This result contrasts findings in the current literature that suggests that positive STRQ is associated with better academic outcomes for students who are academically at-risk (Cadima et al., 2010). For instance, in a study of 191 preschool students with developmental disorders, positive STRQ during preschool was correlated with better learning behaviors, such as persistence, learning strategies, and motivation, during preschool and kindergarten (Rhoad-Drogalis et al., 2018). Additionally, in a study by Allen and colleagues (2013) that investigated if the CLASS measure would predict future student achievement for typically developing secondary students, classrooms with higher STRQ were associated with students with better academic outcomes.

My results may differ from current research due to how STRQ was measured. In the current study, each student's individual STRQ was not captured using the observational class-wide measure of STRQ (i.e., CLASS Emotional Support Domain), rather one STRQ score was provided for the entire classroom. Thus, the CLASS's inability to capture the nuances of STRQ for each student may have influenced the findings, especially as research suggests that STRQ tends to differ between students with and without disabilities (Blacher et al., 2009; McGrath & Bergen, 2015; Prino et al., 2016). Additionally, as discussed above, I used one domain of the CLASS as a proxy for STRQ as the Emotional Support Domain aligned with the STRQ definition in the literature (Rose et al., 2019). This rather unidimensional approach may have missed any subtle differences in STRQ for students with disabilities and delays that a multi-dimensional approach may have captured. Previous research studies have used STRQ measures that are multi-dimensional, such as the Student Teacher Relationship Scale (STRS; Pianta & Steinburg, 1992) which includes thirty items that measure the closeness, conflict, and dependency between students and teachers. While the STRS and CLASS Emotional Support measure similar constructs, the complexity of the STRS may capture nuances related to STRQ that may be missed by the four-item CLASS Emotional Support ratings, which may contribute to the differences in findings between my study and current research trends.

In the present study, the little variance in STRQ across classrooms may have also influenced the non-significant association between STRQ and kindergarten reading outcomes. In a recent study of over 500 students with and without challenges in reading, student-teacher conflict was predictive of lower student reading scores, while student-teacher closeness was not significantly associated with student reading outcomes (Varghese et al., 2019). Thus, because my sample had relatively high STRQ across classrooms, the association between STRQ and student

reading outcomes may align with the results from Varghese and colleague's study of early elementary students (2019), such that high STRQ was not significantly associated with kindergarten reading outcomes.

Finally, my study results may not align with current research trends due to the timing of the data collection. Unlike other studies of STRQ and academic outcomes that are completed across the academic school year (e.g., Hajovsky et al., 2017), my data were collected during Head Start and kindergarten, which was beneficial to learn more about the connection between STRQ in Head Start and kindergarten academic outcomes. In future studies, three data collection points in Head Start related to HS-STC variables (i.e., fall data point), STRQ (i.e., winter data point), and reading outcomes (i.e., spring data point) would be beneficial to examine these constructs across the preschool school year, which aligns with current research (e.g., Zatto & Hoglund, 2019). Thus, collecting data across the same academic year is a possible direction for future research to learn more about the immediate effects of STRQ for preschool students with disabilities and delays.

Limitations

There are multiple limitations related to my dissertation study. First, due to the nature of the study being a secondary data analysis, I relied on the data that were pre-selected by the FACES 2009 researchers. While there are benefits to secondary data analysis, such as alleviating barriers to large-scale data collection and having access to a large sample of students with disabilities, there are also limitations regarding the selection of measures. For example, I was not able to use a broad measure of teacher mental health and examine other related factors, like teacher stress. This is important as the literature suggests that student outcomes are negatively impacted when students are in classrooms with teachers with high levels of stress, poor coping

strategies, and high burnout rates (Herman et al., 2018). Because there were no measures of teacher stress included in the FACES 2009 study, my dissertation relied on teacher self-report of depression using the CES-D as this was the only measure of teacher mental health in the FACES 2009.

If I had included a measure of teacher stress or broader mental health, the study results may have differed. Few teachers in my study reported moderate or severe feelings of depression. Thus, depressive symptoms may not have appropriately characterized teachers' experience. If broader teacher-report measures were included, this may have provided a better opportunity to understand teachers' experience. Teacher stress, burnout, and mental health concerns (e.g., anxiety) are especially salient factors to consider now amidst the rising reports surrounding the COVID-19 pandemic (Liss-Levinson, 2021). The literature suggests that negative student outcomes tend to result when teachers experience high rates of stress and burnout (Herman et al., 2018). Therefore, by only including a teacher report of depressive symptoms, I may have missed additional teacher characteristics that are salient to outcomes for students with disabilities and delays. Thus, researchers are encouraged to consider including measures of teacher mental health, stress, and burnout in future studies. In addition, because the FACES 2009 data are over a decade old, different trends in the data may have emerged overtime, such as the rising rates of teacher mental health difficulties during the COVID-19 pandemic (CDC, 2019b), which may impact the how teacher mental health, STRQ, and academic outcomes intersect.

Another limitation related to secondary data analysis is related to the student behavior variable. After conducting preliminary data analyses, I included a summary score of the externalizing behavioral items in my final model. While this approach aligned with the preliminary data results and previous research (e.g., Hinshaw, 1992), it can also make it

challenging to compare the current findings to other studies that use the FACES 2009 student behavior data, as my study variable will be interpreted differently as I excluded internalizing variables. Thus, a limitation in my study results is the difficulty making comparisons to other FACES 2009 studies that examine student behavior with different samples.

Next, my study was guided by ecological systems theory which supported testing all HS-STC variables in one model. However, due to including so many variables in the primary model, I may have controlled for too many variables at once, which may have made it challenging to see meaningful outcomes related to STRQ as a moderator as well as how each of the HS-STC variables were associated with later academic outcomes. A recommendation for future work is to test different models that were beyond the scope of this study. Based on the current study results, including gender and teacher experience would be important variables to continue to explore, as well as adding broader measures of teacher mental health and stress. Including an informant-report measure for STRQ would also provide additional information about the importance of STRQ for students with disabilities and delays. By including a STRQ measure specific to each student, researchers can investigate the importance of STRQ across students with different abilities (e.g., typically developing, developmental delay, autism, speech and language delays, etc.), which could be beneficial to inform future practice and research.

Using the CLASS as a measure of STRQ was important for this study as I aimed to use an observational measure of STRQ rather than a teacher-self report. The CLASS provided a measure of the overall STRQ amongst all students in the preschool class. However, the CLASS did not provide a specific indicator of the STRQ between each student and their teacher in my sample. Students with disabilities may have had poorer STRQ than their typically developing peers, as commonly suggested by the literature (Blacher et al., 2009; McGrath & Bergen, 2015),

but this was not necessarily captured by the CLASS if the overall STRQ in the class was positive. Similarly, the CLASS did not capture the specific relationships students have with each of the teachers in their classroom, as Head Start classrooms tend to have a primary and secondary teacher. According to the FACES 2009 user guide (USDHHS, 2009a), CLASS observations for the Emotional Support domain included "teachers and students", suggesting that all teachers and students in the classroom may have influenced the CLASS score.

Future directions, as discussed in the following section, should consider how to incorporate both a measure of STRQ at the classroom level using observational methods, while also considering STRQ specific to the student using informant-report measures, such as the Student Teacher Relationship Scale (STRS, Pianta & Steinberg, 1992). This is important as research suggests that STRQ for students with a disability or delay tends to differ from STRQ for typically developing students (Prino et al., 2016). Therefore, understanding each student's personal STRQ could impact how this construct may moderate the association between various HS-STC variables and future academic outcomes for preschool students with disabilities and delays. For example, if students with disabilities experience lower STRQ at baseline, improving STRQ to align with the relationship quality between teachers and typically developing students may be especially influential on future outcomes for students with disabilities and delays. In addition, including an observational measure of STRQ would help further understand how teacher self-reports and third-party observations of STRQ compare for students with disabilities and delays.

Studies should consider using all three CLASS domains (i.e., Instructional Support, Classroom Organization, Emotional Support) to capture all aspects of the classroom environment. Using only the Emotional Support domain as a measure of STRQ makes it

challenging to compare the results of this study to other literature that includes all CLASS domains (e.g., Moen et al., 2019). Thus, while the CLASS Emotional Support Domain tends to be the most strongly correlated with commonly used informant-reports of STRQ (Moen et al., 2019; Walker & Graham, 2021), it can be limiting as it may not capture teacher practices in the classroom that may help facilitate positive STRQ in the classroom, such as quality behavior management skills, student feedback, and language modeling.

Next, I was unable to examine how my findings may have differed by disability group due to my sample size and the measure selected for STRQ. Most students with disabilities and delays included in my study had a speech and language impairment. The small sample sizes across different disability groups in the FACES 2009 impeded me from conducting the moderation analysis with the specific disability groups. Additionally, because my measure of STRQ was at the class level, I was unable to compare the differences in STRQ across groups as the STRQ measure was not a direct measure of each student's relationship with their teacher, rather each student's STRQ in the classroom. Collectively, this made it challenging to examine how STRQ may moderate the association between HS-STC variables and kindergarten reading outcomes for students with differing disability types, as well as compare outcomes for students with and without disabilities.

The results may have differed if I compared outcomes for students of different disability groups. The literature suggests that students with externalizing concerns and neurodevelopmental disorders may experience poorer STRQ compared to students with other disorders, such as Down Syndrome (Eisenhower et al., 2007; Prino et al., 2016). Students with varying disabilities and delays can also experience varying levels of academic performance based on the degree of impairment (Wagner et al., 2006). For instance, students with intellectual disabilities and

multiple disabilities tend to demonstrate poorer academic outcomes compared to students with speech and language disorders. Thus, improvements in STRQ may be especially meaningful for students who experience poorer STRQ and academic outcomes at baseline, as STRQ could serve as a protective factor amidst risk factors.

Due to the use of a latent interaction term in the primary model, I was unable to interpret the absolute model fit using the *Mplus* (Muthén & Muthén, 1998-2017) software. Thus, although the latent interaction added measurement precision to my final model, it also was a study limitation as I was unable to examine how the model fit my current data which consisted of students with disabilities and delays. This is problematic as I was unable to interpret how well my data fit the primary model prior to interpreting the moderation and main effect results. Thus, my model fit was a limitation to the study as it is possible that the latent moderation model may not have been the best representation of the data.

Finally, most Head Start classrooms were observed to have rather high STRQ. Although high STRQ in classrooms is positive, it led to little variability in STRQ for the study. Little variability of STRQ was a limitation as it was challenging to compare differences in STRQ for students with disabilities and delays who had STRQ ratings one standard deviation below the mean (4.83) and one standard deviation above the mean (5.87), as the scores were similar using the seven-point scale. Similar STRQ in classrooms could be attributed to Head Start programming that emphasizes the importance of student-teacher relationships (USDHHS, 2019). Therefore, results may differ in preschool classrooms with less emphasis on cultivating positive STRQ such that the importance of positive STRQ for students with disabilities and delays may be more noticeable when compared to negative STRQ.

Future Directions

Future directions for research and school-based practice are included below as they relate to the results and limitations of my dissertation study.

Research

There is a paucity of work that investigates how STRQ may enhance or buffer preexisting associations between various student, teacher, and classroom variables and student academic outcomes for preschool students with disabilities and delays. Thus, future research should explore alternative models that investigate the moderating role of STRQ. For example, researchers should continue to consider which student, teacher, and classroom variables should be included in their study. Based on the present study findings, salient variables to include in future research are student gender and teacher experience. Additionally, results suggest that Head Start reading outcomes are also important to include when considering kindergarten reading outcomes for students with disabilities and delays.

Additional models should also include both an individual and global measure of STRQ. The Student Teacher Relationship Scale (STRS; Pianta & Steinberg, 1992) is a commonly used individual measure of STRQ in research studies (e.g., e.g., Blacher et al., 2014; Caplan et al., 2016; Hughes, 1999). Aligned with this study, the CLASS Emotional Support Domain can be used as a class wide measure of STRQ. Researchers should examine the magnitude of the association between each STRQ measure and kindergarten reading outcomes to consider the salience of individual and global STRQ for students with disabilities and delays. Additionally, interpreting differences between the moderating role of STRQ using an individual and global measure of STRQ should be examined to learn more about how STRQ at the individual and class level may temper associations between HS-STC variables and kindergarten outcomes for

students with disabilities and delays. Finally, including a model with just an individual STRQ measure (i.e., STRS) would allow researchers to examine the importance of teacher-reported STRQ for students with disabilities and delays, without accounting for class wide STRQ. This approach could be helpful if both STRQ variables are highly correlated when included simultaneously in the model.

An additional direction for studies is to compare models when STRQ is included as an observed variable compared to a latent term. This is an important next step because if the results are comparable, then using an observational STRQ variable may be preferable to decrease model complexity. If model results and trends differed, then limiting measurement error with a latent factor may be preferred. If STRQ were included as an observational variable in the moderation analysis, absolute model fit would be produced, which would be helpful in interpreting how well the data fit the model. This would also address a limitation of the current study as absolute model fit was not interpreted for the primary model due to including the latent interaction. Thus, although it was outside the scope of this dissertation study, multiple models should be examined based on theory and emerging research trends.

When designing my dissertation study, it was challenging to find a data set that included informant-report and observational STRQ measures, the population of interest, and enough data collection time points to answer my research questions. At least three time points were desired to test STRQ's moderating effect between HS-STC variables and kindergarten reading outcomes for students with disabilities and delays (i.e., Time Point 1: HS-STC variables, Time Point 2: STRQ, Time Point 3: Kindergarten Reading Outcomes). The FACES 2009 data used in my study included both the population of interest and enough data collection waves. However, it lacked both a STRQ observational and informant-report measure. The FACES 2014, a more recent

collection of Head Start data, lacked the ideal number of data collection waves and both STRQ measures. The Baby FACES, a similar study completed with children in Early Head Start programs included the population of interest and both STRQ measures but lacked enough data collection waves. Similar trends were found when investigating other large-scale data sets. Thus, to learn more about the moderating effect of STRQ for students with disabilities and delays, large-scale studies that include a nationally representative sample of students with disabilities and delays need to be carefully designed to include these components. For instance, a future study could replicate my study's model using the Baby FACES when enough waves are available to see if similar patterns emerge for younger students with disabilities and delays.

Researchers are also encouraged to think critically about which measures are selected for their study, as well as how variables are measured in their studies. During the preliminary analysis, poor model fit resulted from my confirmatory factor analysis (CFA) for the student behavior variable used in the FACES 2009 study, suggesting that the student behavior items were not all measuring the same construct for my data. However, model fit improved when the items were loaded onto two factors. Thus, before using FACES 2009 data, researchers are encouraged to conduct factor analytic work when using the student behavior variable to examine the structure of the data. More broadly, researchers are advised to complete exploratory and/or confirmatory factor analyses with observed variables to learn more about their measures prior to conducting primary analyses. These preliminary tests help confirm the factor structures of the observed variables in order to examine the construct-related validity. Thus, researchers are encouraged to carefully choose the measures they are using in their study, as well as examine the validity of the measures prior to using them with their selected sample.

Researchers are also encouraged to consider the moderating role of STRQ for students with disabilities and delays between HS-STC variables and other outcomes, such as social-emotional functioning, math outcomes, or academic engagement. This is an important next step based on extant studies that point to the importance of STRQ for multiple student outcomes (McGrath & Van Bergen, 2015). For instance, positive STRQ was associated with higher social emotional outcomes for young students (Decker et al., 2007). Based on the present study findings, researchers should pay special consideration to the importance of STRQ for male students with disabilities and delays, and how this may differ across disability category.

Thus, in addition to including a variety of outcome variables, researchers are encouraged to examine the importance of STRQ for male students across a variety of disabilities and delays. Including students with specific disabilities, such as autism, could provide more information about the role of STRQ for students with unique disabilities. This is an important next step as research suggests that students with varying disabilities can experience differing relationships with their teachers (Prino et al., 2016). Similarly, examining differences related to the importance of STRQ on various school-based outcomes for typically developing students and students with a variety of disabilities would add to the dearth of research available, especially for preschool populations.

School-Based Practice

Teachers and school-based practitioners should consider the importance of STRQ for students with disabilities and delays, particularly for male students. Professionals and teachers working with students should embed strategies to improve STRQ with at-risk populations. The vast STRQ literature suggests that improving closeness between teachers and students, while decreasing conflict fosters positive STRQ (Hughes et al., 2005; Rudasill et al., 2010). Previous

research also supports the connection between positive STRQ and student development (Pianta et al., 2012) and student outcomes (McGrath & Van Bergen, 2015). The present study points to the importance of cultivating positive STRQ, particularly for male students with disabilities and delays.

Teachers can implement strategies to increase closeness and decrease conflict with students. For example, increasing praise and intentionally learning about student interests are ways for teachers to improve STRQ with students (Rimm-Kauman & Sandilos, 2015). Other teacher strategies include spending one-on-one time with students, creating a positive classroom climate, and increasing teacher awareness of how their practices and beliefs influence classroom teaching practices. Educators are also encouraged to collaborate with their colleagues to problem-solve and share strategies related to cultivating positive STRQ.

Additional strategies to improve STRQ that align with the CLASS Emotional Support domain can also be implemented by teachers. Specifically, teachers should aim to improve the positive classroom climate, their sensitive teaching practices, and their regard for student perspectives, while decreasing negative classroom climate. Practically, teachers can foster positive STRQ by improving the emotional tone of the classroom (i.e., positive climate), noticing and responding to student needs (i.e., teacher sensitivity), and considering student interests and perspectives during instruction and while planning lessons (i.e., regard for student perspectives), all while minimizing any hostility or irritability in the classroom (i.e., negative climate); (Li et al., 2020).

Thus, teachers and practitioners are encouraged to prioritize enhancing STRQ for at-risk students, as it can be even more beneficial for these students (Rimm-Kauman & Sandilos, 2015). My study emphasized the importance of STRQ for male students with multiple risk factors,

which aligns with previous research that emphasizes the importance of STRQ for students who are at-risk (Cadima et al., 2010; Fowler et al., 2008). In general, STRQ is associated with beneficial academic, social, and behavioral outcomes for students (Allen et al., 2013; McGrath & Van Bergen, 2015; Pianta & Stuhlman, 2004). Schools, such as Head Start, have recognized the importance of STRQ as they have incorporated relationship building practices into their programming (USDHHS, 2020b). However, it is recommended that more needs to be done for students with disabilities and delays who tend to experience poorer STRQ compared to their typically developing peers (Blacher et al., 2009; McGrath & Bergen, 2015). McNally and Slutsky (2018) argue that students at-risk for developing poor STRQ, like male students with disabilities, would benefit from teachers increasing their understanding of how positive STRQ and student academic outcomes are related. Thus, strategies reported above (e.g., creating a positive classroom environment, spend one-on-one time with students) are suggested for teachers interacting with students presenting with multiple risk factors (Rimm-Kauman & Sandilos, 2015), such as students with disabilities and delays.

A practical approach to enhance STRQ for students with disabilities and delays is to provide teachers with professional developmental opportunities to become more aware of various disabilities and common challenges associated with these diagnoses. After receiving professional development related to special education, general education teachers tend to report more comfort (Sokal & Sharma, 2014) and less concern (Sokal & Sharma, 2017) in teaching students with disabilities. Additionally, teachers report shifts in their attitudes, more acceptance toward students with disabilities, and higher efficacy in teaching students with disabilities when they have received professional development (Crispel & Kasperski, 2021; Sokal & Sharma, 2017). When asked about the impact of professional development, one teacher shared her

perspective by stating, “Before I took the professional development course I would often have conflicts … – now less so; I am more ready to let things slide … I use a lot of self-talk to avoid arguing … Yes, I feel that it has had an effect (Crispel & Kasperski, 2021, pg. 7). Thus, when teachers have a foundational understanding and knowledge about student disabilities, they may be less apt to become irritable when encountering a challenging situation with a student with a disability. Rather, teachers may be more equipped to implement strategies and seek a solution, which can cultivate positive STRQ for students with disabilities and delays.

Specific to male students with disabilities, who from a young age are more likely to experience externalizing concerns (Paz et al., 2020) which can adversely impact STRQ, intentionally fostering STRQ is especially important. Donlevy (2001) suggests that the following four characteristics are “indispensable” to cultivating positive STRQ for teachers and students with disabilities: trust, cooperation, competence, and flexibility. First, trust is arguably a critical element related to STRQ, especially for students with disabilities and delays, as teachers provide ongoing support and feedback amid academic, behavioral, and sometimes emotional challenges. Trust allows students with disabilities the opportunity to learn and grow in a safe environment. Second, teachers of students with disabilities can provide opportunities to model and practice cooperation to bolster STRQ through activities such as sharing, collaborating, and compromising with others. Cooperation is important as it emphasizes the importance of each person in the student-teacher relationship. Third, teachers’ competence while working with students with disabilities is critical to supporting positive STRQ, as certain teacher qualities (i.e., kindness, empathy, patience, teaching amidst challenges, etc.) are especially meaningful to support students with additional needs. Finally, teachers’ flexibility when delivering academic material while not compromising the integrity of the instruction or curricula allows for student interests to

be integrated into daily activities or for the workload to be calibrated to the class's capacity for that day. Ultimately, teachers' flexibility recognizes and responds to the students' needs, which is helpful when fostering positive STRQ as students can feel a sense of support. Trust, cooperation, competence, and flexibility are helpful characteristics to support STRQ for all students with disabilities. However, the present study findings suggest that intentional efforts to bolster STRQ using trust, cooperation, competence, and flexibility are especially important for male students with disabilities and delays due to the positive implications on future academic outcomes.

Similarly, Murray (2002) suggests five recommendations for teachers to enhance STRQ with students with disabilities, which closely align with recommendations for teachers of typically developing students. First, teachers should recognize that students with disabilities want to feel supported by adults in the school. Next, teachers should provide opportunities for students with disabilities to learn how to develop strong relationships with adults, which could be done by teaching students' social emotional skills. Third, teachers are encouraged to get to know their students' values, culture, interests, and family, which is especially important for male students with disabilities from low SES backgrounds as they can be overrepresented in special education categories and differ from teacher demographics, who tend to be predominately female and middle-class (Harry, 1990). Finally, Murray (2002) suggests that teachers should become mindful of interactions in the classroom, as well as teach and set expectations for positive behavior in the classroom. Together, these recommendations encourage teachers to recognize the importance of STRQ, teach and learn from students, as well as model and increase their own awareness while interacting with students with disabilities. And, although Murray's (2002)

article focuses on adolescents with disabilities, these strategies are also applicable for preschool teachers working with male students with disabilities and delays.

Finally, practitioners and researchers are encouraged to develop partnerships to collaboratively develop, implement, and evaluate targeted STRQ evidence-based interventions that can be used in schools to further enhance STRQ between teachers and preschool students with disabilities and delays. While there are numerous strategies and recommendations provided regarding ways to enhance STRQ with students, few evidence-based interventions exist beyond Teacher-Child Interaction Training – Universal (TCIT-U, Gershenson et al., 2010) and the Incredible Years Program (Webster-Stratton, 2005), especially in the context of relationships between teachers and young students with disabilities and delays. Multiple frameworks, such as collaborative action research, demonstrate ways that researchers and early childhood educators can engage in partnerships to support school-based practices (Moran, 2007). Estabrooks and colleagues (2019) outline the Integrated Research-Practice Partnership Process Model, which aims to bridge the research to practice gap by considering how evidence-based interventions can be applied in practice with the target student population. This is done by considering sustainability, practicality, and adaptations of evidence-based practices to meet the needs of schools and students. Ultimately, this framework is effective due to the collaborative partnership between researchers who have expertise in evidence-based practices and educators who have expertise in school-based practices and their student needs. Thus, practitioners and researchers can jointly propel this work forward, especially for the most vulnerable youth.

Summary and Conclusion

Positive student-teacher relationship quality (STRQ) is associated with beneficial academic outcomes for school-aged typically developing students (Hamre & Pianta, 2001). Yet,

there is a dearth of studies that investigate the intricacies of STRQ for preschool students with disabilities and delays, despite evidence that STRQ may be especially meaningful for this group (McGrath & Van Bergen, 2015). Using secondary data from the Family and Child Experiences Survey (FACES) 2009 data set, I aimed to address this gap by investigating the following three aims: 1) the association between STRQ in Head Start and kindergarten reading outcomes; 2) the association between Head Start student, teacher, classroom (HS-STC) variables and kindergarten reading outcomes; and 3) the moderating role of STRQ between (HS-STC) variables and kindergarten reading outcomes. This study incorporated an observational measure of STRQ using the Emotional Support domain from the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) to reduce rater biases. All research questions were tested via one structural equation model that included STRQ as a latent term to decrease measurement error.

The results suggested that STRQ moderated the association between student gender and kindergarten reading outcomes. Specifically, males' kindergarten reading outcomes improved when STRQ was high during preschool. STRQ did not buffer or enhance the association between other Head Start student, teacher, and classroom variables and kindergarten academic outcomes for students with disabilities and delays. The findings also suggested that higher kindergarten reading outcomes were experienced by students with disabilities and delays who had Head Start teachers with less experience and for students with higher Head Start reading outcomes. Other associations between Head Start student, teacher, and classroom variables and kindergarten academic outcomes for students with disabilities and delays were not significant. Finally, STRQ in Head Start was not associated with kindergarten reading outcomes for students with disabilities and delays.

Overall, this study contributed to the lack of research related to understanding the importance of STRQ for young students with disabilities and delays. Additional research is needed to further investigate the nuances of STRQ for students with disabilities and delays, especially for male students. Alternative models beyond the scope of this study, such as those that include various measures and samples, should be tested. Implications for practice include embedding strategies within classrooms to foster positive STRQ, as well as cultivating research-practitioner collaborations to propel this work forward in efforts to enhance outcomes for young populations with disabilities and delays.

REFERENCES

Acar, I. H., Veziroğlu-Çelik, M., Rudasill, K. M., & Sealy, M. A. (2022, February). Preschool children's self-regulation and learning behaviors: The moderating role of teacher–child relationship. In *Child & Youth Care Forum* (Vol. 51, No. 1, pp. 1-18). Springer US.

Adani, S., & Cepanec, M. (2019). Sex differences in early communication development: behavioral and neurobiological indicators of more vulnerable communication system development in boys. *Croatian medical journal*, *60*(2), 141-149.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. *School psychology review*, *42*(1), 76-98.

American Academy of Pediatrics Council on Community Pediatrics. Poverty and child health in the United States. *Pediatrics.* https://doi.org/10.1542/peds.2016-0339

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Psychological Association. (2022). Effects of Poverty, Hunger and Homelessness on Children and Youth. https://www.apa.org/pi/families/poverty

Ang, R. P. (2005). *Teacher–Student relationship inventory* doi: http://dx.doi.org/10.1037/t08044 000

Ang, R. P., Ong, S. L., & Li, X. (2020). Student version of the Teacher–Student Relationship Inventory (S-TSRI): Development, validation and invariance. *Frontiers in Psychology*, *11*, 1724.

Ansari, A., Hofkens, T. L., & Pianta, R. C. (2020). Teacher-student relationships across the first seven years of education and adolescent outcomes. *Journal of Applied Developmental Psychology*, *71*, 101200.

Asparouhov, T. & Muthén, B. (2021). First-Order Derivative Warning Message, Condition Number, and Non-Identification. https://statmodel.com/download/ConditionNumber.pdf

Baker, C. N., Peele, H., Daniels, M., Saybe, M., Whalen, K., Overstreet, S., & The New Orleans, T. I. S. L. C. (2021). The experience of COVID-19 and its impact on teachers' mental Health, coping, and teaching. *School Psychology Review*, *50*(4), 491-504.

Banerjee, N., Stearns, E., Moller, S., & Mickelson, R. A. (2017). Teacher job satisfaction and student achievement: The roles of teacher professional community and teacher collaboration in schools. *American Journal of Education*, *123*(2), 000-000.

Bates, L. A., & Glick, J. E. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Social science research*, *42*(5), 1180-1190.

Benischek, A., Long, X., Rohr, C. S., Bray, S., Dewey, D., & Lebel, C. (2020). Pre-reading language abilities and the brain's functional reading network in young children. *NeuroImage*, *217*, 116903.

Bentler (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, *88*(3), 588.

Berliner, D. (2001). Learning about and learning from expert teachers. International Journal of Educational Research 35 (5): 463-482.

Berliner, D. (2004). Teacher expertise. In Teaching and learning in the secondary school, edited by Ann Shelton Mayes and Bob Moon, pp. 107-113. New York, NY: Routledge Falmer.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. *Journal of school psychology*, *46*(2), 151-172.

Berry, D. J., Bridges, L. J., & Zaslow, M. J. (2004). Early childhood measures profiles. Washington, DC: Child Trends.

Blacher, J., Baker, B. L., & Eisenhower, A. S. (2009). Student–teacher relationship stability across early school years for children with intellectual disability or typical development. American journal on intellectual and developmental disabilities, 114(5), 322-339.

Blacher, J., Howell, E., Lauderdale-Littin, S., Reed, F. D. D., & Laugeson, E. A. (2014). Autism spectrum disorder and the student teacher relationship: A comparison study with peers with intellectual disability and typical development. Research in Autism Spectrum Disorders, 8(3), 324-333.

Bronfenbrenner, U. (1979). The Ecology of Human Development. Cambridge, Mass: Harvard University Press.

Bronfenbrenner, U. (1994). Ecological models of human development. In T. Husen & T. N. Postlewaite (Eds.), *International Encyclopedia of Education* (2nd ed., Vol. 3, pp. 1643-1647), Oxford, UK: Pergamon Press.

Briones, S. K. F., Dagamac, R. J. R., David, J. D., & Landerio, C. A. B. (2021). Factors

Affecting the Students' Scholastic Performance: A Survey Study. *Indonesian Journal of Educational Research and Technology*, *2*(2), 97-102.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In: K. A. Bollen & J. S. Long (Eds.), Testing Structural Equation Models (pp. 136–162). Beverly Hills, CA: Sage

Brunsting, N. C., Sreckovic, M. A., & Lane, K. L. (2014). Special education teacher burnout: A synthesis of research from 1979 to 2013. *Education and treatment of children*, 681-711.

Bulotsky-Shearer, R. J., Fernandez, V., Dominguez, X., & Rouse, H. L. (2011). Behavior problems in learning activities and social interactions in Head Start classrooms and early reading, mathematics, and approaches to learning. *School Psychology Review*, *40*(1), 39-56.

Burchinal, M. R., Peisner-Feinberg, E., Pianta, R., & Howes, C. (2002). Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of school psychology*, *40*(5), 415-436.

Burdick-Will, J.,, Ludwig, J.,, Raudenbush, S.W.,, Sampson, R.J.,, Sanbonmatsu, L., , & Sharkey, P. (2011). "Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-Experimental, and Observational Comparison." In Whither Opportunity?, edited by Greg J. Duncan and Richard J. Murnane: Russell Sage Foundation: 255-276.

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

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher–student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of school psychology*, *48*(6), 457-482.

Caemmerer, J. M., & Keith, T. Z. (2015). Longitudinal, reciprocal effects of social skills and achievement from kindergarten to eighth grade. *Journal of school psychology*, *53*(4), 265-281.

Camacho, K. A., & Krezmien, M. P. (2020). A statewide analysis of school discipline policies and suspension practices. *Preventing School Failure: Alternative Education for Children and Youth*, *64*(1), 55-66.

Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. Development and Psychopathology, 12, 467– 488

Caplan, B., Feldman, M., Eisenhower, A., & Blacher, J. (2016). Student–teacher relationships for

young children with autism spectrum disorder: Risk and protective factors. *Journal of autism and developmental disorders*, *46*(12), 3653-3666.

Casillas, A., Robbins, S., Allen, J., Kuo, Y. L., Hanson, M. A., & Schmeiser, C. (2012). Predicting early academic failure in high school from prior academic achievement, psychosocial characteristics, and behavior. *Journal of Educational Psychology*, *104*(2), 407.

Castro-Ramirez, F., Al-Suwaidi, M., Garcia, P., Rankin, O., Ricard, J. R., & Nock, M. K. (2021). Racism and poverty are barriers to the treatment of youth mental health concerns. *Journal of Clinical Child & Adolescent Psychology*, *50*(4), 534-546.

Centers for Disease Control and Prevention. (2019a). Autism Spectrum Disorder (ASD). https://www.cdc.gov/ncbddd/autism/addm-community-report/delay-to-accessing-services.html#:~:text=A%20medical%20diagnosis%20can%20be,4%20years%20and%204%20months).

Centers for Disease Control and Prevention. (2019b). Mental Health Impact of the COVID-19 Pandemic on Teachers and Parents of K-12 Students. https://www.cdcfoundation.org/mental-health-triangulated-report?inline

Center for Disease Control and Prevention. (2020). Disability and Health Overview. https://www.cdc.gov/ncbddd/disabilityandhealth/disability.html

Center on PBIS. (2022). Positive Behavioral Interventions & Supports [Website]. www.pbis.org.

Center on the Social and Emotional Foundations for Early Learning (CSEFEL). 2022. Resources: Preschool Training Modules. http://csefel.vanderbilt.edu/resources/training_preschool.html

Chen, S., & Phillips, B. (2018). Exploring teacher factors that influence teacher-child relationships in Head Start: A grounded theory. *The Qualitative Report*, *23*(1), 80-97

Chen, S., Phillips, B., & Izci, B. (2018). Teacher–child relational conflict in Head Start exploring the roles of child behaviour, teacher stress, and bias, and classroom environment. *Early Child Development and Care*..

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.

Claessens, A., Engel, M., & Curran, F. C. (2014). Academic content, student learning, and the persistence of preschool effects. *American Educational Research Journal*, *51*(2), 403-434.

Clarke, P. (2008). When can group level clustering be ignored? Multilevel models versus single

level models with sparse data. *Journal of Epidemiology & Community Health*, *62*(8), 752-758.

Cooc, N. (2017). Examining racial disparities in teacher perceptions of student disabilities. *Teachers College Record*, *119*(7), 1-32.

Connor, C. M., Son, S. H., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, *43*(4), 343-375.

Crispel, O., & Kasperski, R. (2021). The impact of teacher training in special education on the implementation of inclusion in mainstream classrooms. *International Journal of Inclusive Education*, *25*(9), 1079-1090.

Croninger, R. G., Rice, J. K., Rathbun, A., & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of education review*, *26*(3), 312-324.

Cruz, R. A., & Rodl, J. E. (2018). An integrative synthesis of literature on disproportionality in special education. *The Journal of Special Education*, *52*(1), 50-63.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. *Learning Policy Institute*.

Davis, H. A. (2001). *Quality of the Student/Teacher relationship scale* doi: http://dx.doi.org/10.1037/t21888-000

Davis, H. A., & Lease, A. M. (2007). Perceived organizational structure for teacher liking: The role of peers' perceptions of teacher liking in teacher–student relationship quality, motivation, and achievement. *Social Psychology of Education*, *10*(4), 403-427.

Decker, D. M., Dona, D. P., & Christenson, S. L. (2007). Behaviorally at-risk African American students: The importance of student–teacher relationships for student outcomes. *Journal of School Psychology*, *45*(1), 83-109.

Demirkaya, P. N., & Bakkaloglu, H. (2015). Examining the student-teacher relationships of children both with and without special needs in preschool classrooms. Educational Sciences: Theory & Practice, 15(1).

Domínguez Escalo´n, X., & Greenfield, D. B. (2009). Learning behaviors mediating the relationship between behavior problems and academic outcomes. NHSA Dialog, 12, 1–17.

Domitrovich, C. E., Morgan, N. R., Moore, J. E., Cooper, B. R., Shah, H. K., Jacobson, L., &

Greeenberg, M. T. (2013). One versus two years: Does length of exposure to an enhanced preschool program impact the academic functioning of disadvantaged children in kindergarten?. *Early Childhood Research Quarterly*, *28*(4), 704-713.

Donlevy, J. (2001). High-stakes environments and effective student-teacher relationships: Some lessons from special education. *International Journal of instructional media*, *28*(1), 1.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43*(6), 1428-1446.

Durkin, M. S., Maenner, M. J., Meaney, F. J., Levy, S. E., DiGuiseppi, C., Nicholas, J. S., ... & Schieve, L. A. (2010). Socioeconomic inequality in the prevalence of autism spectrum disorder: evidence from a US cross-sectional study. *PloS one*, *5*(7), e11551.

Eisenhower, A. S., Baker, B. L., & Blacher, J. (2007). Early student–teacher relationships of children with and without intellectual disability: Contributions of behavioral, social, and self-regulatory competence. *Journal of School Psychology*, *45*(4), 363-383.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. Structural equation modeling, 8(3), 430-457.

Estabrooks, P. A., Harden, S. M., Almeida, F. A., Hill, J. L., Johnson, S. B., Porter, G. C., & Greenawald, M. H. (2019). Using integrated research-practice partnerships to move evidence-based principles into practice. *Exercise and sport sciences reviews*, *47*(3), 176.

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.

Eyberg. S. (1988). Parent child interaction therapy: integration of traditional and behavioral concerns. Child Fam Behav Therapy 10(1):3346.

Filges, T., Sonne-Schmidt, C. S., & Nielsen, B. C. V. (2018). Small class sizes for improving student achievement in primary and secondary schools: a systematic review. *Campbell Systematic Reviews*, *14*(1), 1-107.

Folmer-Annevelink, E., Doolaard, S., Mascareño, M., & Bosker, R. J. (2010). Class size effects on the number and types of student-teacher interactions in primary classrooms. *The Journal of Classroom Interaction*, 30-38.

Fowler, L. T. S., Banks, T. I., Anhalt, K., Der, H. H., & Kalis, T. (2008). The association between externalizing behavior problems, teacher-student relationship quality, and academic performance in young urban learners. *Behavioral Disorders*, *33*(3), 167-183.

Francis, J., & Barnett, W. S. (2019). Relating preschool class size to classroom quality and student achievement. *Early Childhood Research Quarterly*, *49*, 49-58.

Freire, S., Pipa, J., Aguiar, C., Vaz da Silva, F., & Moreira, S. (2020). Student–teacher closeness and conflict in students with and without special educational needs. *British Educational Research Journal*, *46*(3), 480-499.

Garson, G. D. (2012). Testing statistical assumptions. *Asheboro, NC: Statistical Associates Publishing*.

Gershenson, R., Lyon, A., & Budd, K. S. (2010). Promoting positive interactions in the classroom: Adapting Parent-Child Interaction Therapy as a universal prevention program. *Education and Treatment of Children*, *33,* 261-287.

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions. *Yale University Child Study Center*, *9*(28), 1-16.

Gilliam, W. S., & Shahar, G. (2006). Preschool and child care expulsion and suspension: Rates and predictors in one state. *Infants & Young Children*, *19*(3), 228-245.

Gilmour, A. F., Fuchs, D., & Wehby, J. H. (2019). Are students with disabilities accessing the curriculum? A meta-analysis of the reading achievement gap between students with and without disabilities. *Exceptional Children*, *85*(3), 329-346.

Glazzard, J., & Rose, A. (2019). The impact of teacher well-being and mental health on pupil progress in primary schools. *Journal of Public Mental Health*.

Goldbach, J. T., & Gibbs, J. J. (2017). A developmentally informed adaptation of minority stress for sexual minority adolescents. *Journal of adolescence*, *55*, 36-50.

Goldie, P. D., & O'Connor, E. E. (2021). The Gender Achievement Gap: Do Teacher-Student Relationships Matter?. *Psi Chi Journal of Psychological Research*, *26*(2).

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of school psychology*, *45*(1), 3-19.

Griggs, M. S., Gagnon, S. G., Huelsman, T. J., Kidder-Ashley, P., & Ballard, M. (2009). Student–teacher relationships matter: Moderating influences between temperament and preschool social competence. *Psychology in the Schools*, *46*(6), 553-567.

GSRP Implementation Manual. (2020). Classroom Requirements. https://www.michigan.gov/documents/mde/Classroom_Requirements._August_2014_466777_7.pdf

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Multivariate Data Analysis fifth edition Prentice-Hall. *Upper Saddle River*.

Hajovsky, D. B., Mason, B. A., & McCune, L. A. (2017). Teacher-student relationship quality and academic achievement in elementary school: A longitudinal examination of gender differences. *Journal of school psychology*, *63*, 119-133.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child development*, *72*(2), 625-638.

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

Hamre, B. K., & Pianta, R. C. (2004). Self-reported depression in nonfamilial caregivers: Prevalence and associations with caregiver behavior in child-care settings. *Early childhood research quarterly*, *19*(2), 297-318.

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. Doi:10.1086/669616.

Hanushek, E. A., Peterson, P. E., Talpey, L. M., & Woessmann, L. (2019). *The unwavering SES achievement gap: Trends in US student performance* (No. w25648). National Bureau of Economic Research.

Harding, S., Morris, R., Gunnell, D., Ford, T., Hollingworth, W., Tilling, K., ... & Kidger, J. (2019). Is teachers' mental health and wellbeing associated with students' mental health and wellbeing?. *Journal of affective disorders*, *242*, 180-187.

Harry, B., & Klingner, J. (2014). *Why are so many minority students in special education?*. Teachers College Press.

Hauser-Cram, P., Durand, T. M., & Warfield, M. E. (2007). Early feelings about school and later academic outcomes of children with special needs living in poverty. *Early Childhood Research Quarterly*, *22*(2), 161-172.

Hauser-Cram, P., & Woodman, A. C. (2016). Trajectories of internalizing and externalizing behavior problems in children with developmental disabilities. *Journal of abnormal child psychology*, *44*(4), 811-821.

Herman, K. C., Hickmon-Rosa, J. E., & Reinke, W. M. (2018). Empirically derived profiles of teacher stress, burnout, self-efficacy, and coping and associated student outcomes. *Journal of Positive Behavior Interventions*, *20*(2), 90-100.

Hindman, A. H., & Bustamante, A. S. (2019). Teacher depression as a dynamic variable: Exploring the nature and predictors of change over the head start year. *Journal of applied developmental psychology*, *61*, 43-55.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: causal relationships and underlying mechanisms. *Psychological bulletin*, *111*(1), 127.

Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of clinical child psychology*, *21*(2), 143-150.

Hirsch, S. E., Driver, M. K., Hinzman-Ferris, M., & Bruhn, A. (2022). Identifying Students with Intensive Academic and Behavioral Needs. In *Delivering Intensive, Individualized Interventions to Children and Youth with Learning and Behavioral Disabilities* (Vol. 32, pp. 7-28). Emerald Publishing Limited.

Hoffman, S. (2018). *The Effects of School Socio-Economic Status on Oral Reading Fluency Scores of Students Identified and Not Identified with a Specific Learning Disability*. Indiana University of Pennsylvania.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods,* 3(4), 424-453

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Huan, V. S. L., Quek, C. L., Yeo, L. S., Ang, R. P., & Chong, W. H. (2012). How teacher student relationship influenced student attitude towards teachers and school.

Hughes, J. N., Gleason, K. A., & Zhang, D. (2005). Relationship influences on teachers' perceptions of academic competence in academically at-risk minority and majority first grade students. *Journal of school psychology*, *43*(4), 303-320.

Hynek, K. A., Abebe, D. S., Hollander, A. C., Liefbroer, A. C., Hauge, L. J., & Straiton, M. L. (2022). The association between persistent low parental income during preschool age and mental disorder in adolescence and early adulthood: a Norwegian register-based study of migrants and non-migrants. *BMC psychiatry*, *22*(1), 1-9.

IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.

Individuals With Disabilities Education Act, 3 U.S.C. § 1401 (2004).

Jackson, D. L., Gillaspy Jr, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological methods*, *14*(1), 6.

Jeon, L., Buettner, C. K., & Grant, A. A. (2018). Early childhood teachers' psychological well being: Exploring potential predictors of depression, stress, and emotional exhaustion. *Early education and development*, *29*(1), 53-69.

Jeon, S., Jeon, L., Lang, S., & Newell, K. (2021). Teacher Depressive Symptoms and Child Math Achievement in Head Start: The Roles of Family–Teacher Relationships and Approaches to Learning. *Child development*.

Jerome, E. M., Hamre, B. K., & Pianta, R. C. (2009). Teacher–child relationships from kindergarten to sixth grade: Early childhood predictors of teacher-perceived conflict and closeness. *Social development*, *18*(4), 915-945.

Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories: Factors associated with change. Journal of Educational Psychology, 91(1), 116 126. https://doi.org/10.1037/0022-0663.91.1.116

Johnston, R.S., & J. Watson. (2005). The effects of synthetic phonics teaching on reading and spelling attainment, a seven year longitudinal study. Scottish Executive Education Department. http://www.scotland.gov.uk/Resource/Doc/36496/0023582.pdf

Justice, L. M., Logan, J. A., Lin, T. J., & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science*, *25*(9), 1722-1729.

Kastner, J. W., May, W., & Hildman, L. (2001). Relationship between language skills and academic achievement in first grade. *Perceptual and motor skills*, *92*(2), 381-390.

Kenny, D. A. (2020). Measuring Model Fit. http://www.davidakenny.net/cm/fit.htm

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Koenen, K. A. B., De Vroey, A., Kelchtermans, G., & Spilt, J. L. (2021). Strengthening Individual Teacher-Child Relationships: An Intervention Study Among Student Teachers in Special Education. *Front. Educ. 6: 769573. doi: 10.3389/feduc*.

Kremer, K. P., Flower, A., Huang, J., & Vaughn, M. G. (2016). Behavior problems and children's academic achievement: A test of growth-curve models with gender and racial differences. *Children and youth services review*, *67*, 95-104.

Krezmien, M. P., Leone, P. E., & Achilles, G. M. (2006). Suspension, race, and disability:

Analysis of statewide practices and reporting. *Journal of emotional and behavioral disorders*, *14*(4), 217-226.

Kulkarni, T., Sullivan, A. L., & Kim, J. (2020). Externalizing behavior problems and low academic achievement: Does a causal relation exist?. *Educational Psychology Review*, 1-22.

Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy*, *12*(2), 241-279.

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, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity?. *Educational researcher*, *31*(1), 3-12.

Lee, J. S. (2012). The effects of the teacher–student relationship and academic press on student engagement and academic performance. *International Journal of Educational Research*, *53*, 330-340.

Lee, K., & Luellen, J. L. (2021). Long-Term Head Start Impacts on Cognitive Outcomes for Children With Disabilities. *Families in Society*, *102*(1), 104-118.

Lee, K., & Rispoli, K. (2016). Effects of individualized education programs on cognitive outcomes for children with disabilities in head start programs. *Journal of Social Service Research*, *42*(4), 533-547.

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: an international journal*, *46*(3), 517-528.

Li, H., Liu, J., & Hunter, C. V. (2020). A Meta-Analysis of the Factor Structure of the Classroom Assessment Scoring System (CLASS). *The Journal of Experimental Education*, *88*(2), 265-287.

Liederman, J., Kantrowitz, L., & Flannery, K. (2005). Male vulnerability to reading disability is not likely to be a myth: A call for new data. *Journal of learning disabilities*, *38*(2), 109-129.

Liew, J., Chen, Q., & Hughes, J. N. (2010). Child effortful control, teacher–student relationships, and achievement in academically at-risk children: Additive and interactive effects. *Early Childhood Research Quarterly*, *25*(1), 51-64.

Liss-Levinson, A. (2021). K-12 public school employee views on finances, employment outlook,

and safety concerns due to COVID-19. *Mission Square Research Institute.* https://www.slge.org/wp-content/uploads/2021/09/k-12surveyfindings_sept21.pdf

Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational review*, *62*(2), 175-187.

Loomis, A. M. (2021). The influence of early adversity on self-regulation and student-teacher relationships in preschool. *Early Childhood Research Quarterly*, *54*, 294-306.

Loomis, A. M., Freed, S., & Coffey, R. (2022). Inhibitory Control, Student–Teacher Relationships, and Expulsion Risk in Preschools: An Indirect Effects Path Analysis. *Early Childhood Education Journal*, 1-10.

MacArdy, A. L. (2008). Jamie S. v. Milwaukee public schools: Urban challenges cause systemic violations of the IDEA. *Marq. L. Rev.*, *92*, 857.

Mandlawitz, M. (2003). A Tale of 3 Cities: Urban Perspectives on Special Education.

Mantzicopoulos, P. (2005). Conflictual relationships between kindergarten children and their teachers: Associations with child and classroom context variables. *Journal of School Psychology*, *43*(5), 425-442.

Manu, M., Torppa, M., Eklund, K., Poikkeus, A. M., Lerkkanen, M. K., & Niemi, P. (2021). Kindergarten pre-reading skills predict Grade 9 reading comprehension (PISA Reading) but fail to explain gender difference. *Reading and Writing*, *34*(3), 753-771.

Martin, A. J., Strnadová, I., Němec, Z., Hájková, V., & Květoňová, L. (2021). Teacher assistants working with students with disability: the role of adaptability in enhancing their workplace wellbeing. *International Journal of Inclusive Education*, *25*(5), 565-587.

Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. *Journal of psychoeducational assessment*, *24*(4), 367-380.

Mawene, D., & Bal, A. (2018). Factors Influencing Parents' Selection of Schools for Children with Disabilities: A Systematic Review of the Literature. *International Journal of Special Education*, *33*(2), 313-329.

McGrath, K. F., & Van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student–teacher relationships and their outcomes. *Educational Research Review*, *14*, 1-17.

McGrew, K. S. & Woodcock, R. W. (2001). Woodcock-Johnson III Technical Manual. Itasca, IL: Riverside Publishing.

McLean, L., & Connor, C. M. (2015). Depressive symptoms in third-grade teachers: Relations to

classroom quality and student achievement. *Child development*, *86*(3), 945-954.

McNally, S., & Slutsky, R. (2018). Teacher–child relationships make all the difference: Constructing quality interactions in early childhood settings. *Early Child Development and Care*, *188*(5), 508-523.

Memon, M. A., Cheah, J. H., Ramayah, T., Ting, H., Chuah, F., & Cham, T. H. (2019). Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, *3*(1), 1-11.

Miller-Lewis, L. R., Sawyer, A. C., Searle, A. K., Mittinty, M. N., Sawyer, M. G., & Lynch, J. . (2014). Student-teacher relationship trajectories and mental health problems in young children. *BMC psychology*, *2*(1), 1-18.

Moen, A. L., Sheridan, S. M., Schumacher, R. E., & Cheng, K. C. (2019). Early Childhood Student–Teacher Relationships: What is the Role of Classroom Climate for Children Who are Disadvantaged?. *Early Childhood Education Journal*, *47*(3), 331-341.

Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate behavioral research*, *39*(1), 129-149.

Monahan, B. (2020). The Ways Teacher-Student Relationships in an Independent Study Setting Influence Student Achievement.

Mora-Ruano, J. G., Heine, J. H., & Gebhardt, M. (2019). Does teacher collaboration improve student achievement? analysis of the german PISA 2012 sample. In *Frontiers in Education* (Vol. 4, p. 85). Frontiers

Moran, M. J. (2007). Collaborative action research and project work: Promising practices for developing collaborative inquiry among early childhood preservice teachers. *Teaching and Teacher Education*, *23*(4), 418-431.

Murray, C. (2002). Supportive teacher-student relationships: Promoting the social and emotional health of early adolescents with high incidence disabilities. *Childhood Education*, *78*(5), 285-290.

Murray, C., Murray, K. M., & Waas, G. A. (2008). Child and teacher reports of teacher–student relationships: Concordance of perspectives and associations with school adjustment in urban kindergarten classrooms. *Journal of applied developmental psychology*, *29*(1), 49-61.

Murray, C., & Pianta, R. C. (2007). The importance of teacher-student relationships for adolescents with high incidence disabilities. *Theory into practice*, *46*(2), 105-112.

Muthén, L. K., & Muthén, B. O. (1998 –2021). Mplus user's guide (8th ed.). Los Angeles, CA: Author.

Nalipay, M. J. N., Cai, Y., & King, R. B. (2020). Why do girls do better in reading than boys? How parental emotional contagion explains gender differences in reading achievement. *Psychology in the Schools*, *57*(2), 310-319.

National Center for Education Statistics. (2006). Facts from NLTS2: School Behavior and Disciplinary Experiences of Youth with Disabilities. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=NCSER2006NLTS2F

National Center for Education Statistics. (2021). Students with Disabilities. https://nces.ed.gov/programs/coe/indicator/cgg

National Center for Learning Disabilities. (2017). Supporting Academic Success. https://www.ncld.org/research/state-of-learning-disabilities/supporting-academic-success/

National Center for Learning Disabilities. (2020). Significant Disproportionality in Special Education: The Role of Income. https://www.ncld.org/wp-content/uploads/2020/10/2020-NCLD-Disproportionality_-Low-Income-Students_FINAL.pdf

National Center for Pyramid Model Innovations (NCPMI). (2022). The Pyramid Model. https://challengingbehavior.cbcs.usf.edu/Pyramid/overview/index.html

National Center for Special Education Research. (2006). Facts from NLTS2: School Behavior and Disciplinary Experiences of Youth with Disabilities. https://ies.ed.gov/ncser/pubs/nlts2200603/disability.asp

National Institute of Health. (2020). Report of the National Reading Panel. https://www.nichd.nih.gov/publications/pubs/nrp/findings

Nelson, C. A. & Sheridan, M.A. (2011). "Lessons from Neuroscience Research for Understanding Causal Links Between Family and Neighborhood Characteristics and Educational Outcomes." In Whither Opportunity?, edited by Greg J. Duncan and Richard J. Murnane: Russell Sage Foundation: 27-46.

O'Connor, E. E., Dearing, E., & Collins, B. A. (2011). Teacher-child relationship and behavior problem trajectories in elementary school. *American Educational Research Journal*, *48*(1), 120-162.

O'Connor, E., & McCartney, K. (2007). Examining teacher–child relationships and achievement as part of an ecological model of development. *American educational research journal*, *44*(2), 340-369.

Odom, S. L. (2000). Preschool inclusion: What we know and where we go from here. *Topics in early childhood special education*, *20*(1), 20-27.

Okano, L., Jeon, L., Crandall, A., Powell, T., & Riley, A. (2020). The cascading effects of

externalizing behaviors and academic achievement across developmental transitions: Implications for prevention and intervention. *Prevention Science*, *21*(2), 211-221.

Olivier, E., Morin, A. J., Langlois, J., Tardif-Grenier, K., & Archambault, I. (2020). Internalizing and externalizing behavior problems and student engagement in elementary and secondary school students. *Journal of youth and adolescence*, *49*(11), 2327-2346.

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: implications for early screening and theories of reading. *Developmental science*, *20*(5), e12471.

Patterson, J. H., Collins, L., & Abbott, G. (2004). A Study of Teacher Resilience in Urban Schools. *Journal of Instructional Psychology*, *31*(1).

Paz, Y., Orlitsky, T., Roth-Hanania, R., Zahn-Waxler, C., & Davidov, M. (2021). Predicting externalizing behavior in toddlerhood from early individual differences in empathy. *Journal of child psychology and psychiatry*, *62*(1), 66-74.

Peng, P., Zhang, Z., Wang, W., Lee, K., Wang, T., Wang, C., ... & Lin, J. (2022). A meta analytic review of cognition and reading difficulties: Individual differences, moderation, and language mediation mechanisms. *Psychological Bulletin*, *148*(3-4), 227.

Peterson, J. L., & Zill, N. (1986). Marital disruption, parent-child relationships, and behavior problems in children. *Journal of Marriage and the Family*, 295-307.

Pianta, R. C. (2001). Student-Teacher Relationship Scale Professional Manual. Psychological Assessment Resources, Inc. https://education.virginia.edu/sites/default/files/uploads/resourceLibrary/STRS_Professional_Manual.pdf

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In *Handbook of research on student engagement* (pp. 365-386). Springer, Boston, MA.

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

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). Classroom Assessment Scoring System™: Manual K-3. Paul H Brookes Publishing.

Pianta, R. C., & Steinberg, M. (1992). Teacher–child relationships and the process of adjusting to school.

Pianta, R. C., & Stuhlman, M. W. (2004). Teacher-child relationships and children's success in the first years of school. *School psychology review*, *33*(3), 444-458.

Podolsky, A., Kini, T., & Darling-Hammond, L. (2019). Does teaching experience increase teacher effectiveness? A review of US research. *Journal of Professional Capital and Community*.

Prino, L. E., Pasta, T., Gastaldi, F. G. M., & Longobardi, C. (2016). The Effect of Autism Spectrum Disorders, Down Syndrome, Specific Learning Disorders and Hyperactivity and Attention Deficits on the Student-Teacher Relationship. Electronic Journal of Research in Educational Psychology, 14(1), 89-106.

Radloff, L. S. (1977). The CES-D scale: A self report depression scale for research in the general population. Applied Psychological Measurements, 1, 385-401.

Ramsook, K. A., Welsh, J. A., & Bierman, K. L. (2020). What you say, and how you say it: Preschoolers' growth in vocabulary and communication skills differentially predict kindergarten academic achievement and self-regulation. *Social Development*, *29*(3), 783-800.

Reardon, S. F., Fahle, E. M., Kalogrides, D., Podolsky, A., & Zárate, R. C. (2019). Gender achievement gaps in US school districts. *American Educational Research Journal*, *56*(6), 2474-2508.

Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education*, *38*(3), 130-143.

Reynolds, C. R., & Kamphaus, R. W. (2015). Behavior Assessment Scoring System Third Edition. *Pearson.* Personassessments.com.

Rhoad-Drogalis, A., Justice, L. M., Sawyer, B. E., & O'Connell, A. A. (2018). Teacher–child relationships and classroom-learning behaviours of children with developmental language disorders. *International journal of language & communication disorders*, *53*(2), 324-338.).

Richards, C. M., Symons, D. K., Greene, C. A., & Szuszkiewicz, T. A. (1995). The bidirectional relationship between achievement and externalizing behavior problems of students with learning disabilities. *Journal of learning disabilities*, *28*(1), 8-17.

Rimm-Kauman, S., & Sandilos, L. (2015). Improving Students' Relationships with Teachers to Provide Essential Supports for Learning. American Psychological Association. https://www.apa.org/education-career/k12/relationships

Roorda, D. L., Koomen, H. M., Spilt, J. L., & Oort, F. J. (2011). The influence of affective

teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of educational research*, *81*(4), 493-529.

Rose, J., Leverson, M., & Smith, K. (2019). PBIS Forum 18: Practice Brief Cultivating Positive Student-Teacher Relationships. PBIS Leadership Forum. https://assets-global.website-files.com/5d3725188825e071f1670246/5d7bd214b509ea64bc58ac5a_rdq7%20brief%20-%20student-teacher%20relationships-2.pdf

Rudasill, K. M., Reio Jr, T. G., Stipanovic, N., & Taylor, J. E. (2010). A longitudinal study of student–teacher relationship quality, difficult temperament, and risky behavior from childhood to early adolescence. *Journal of school psychology*, *48*(5), 389-412.

Rudasill, K. M., Rimm-Kaufman, S. E., Justice, L. M., & Pence, K. (2006). Temperament and language skills as predictors of teacher–child relationship quality in preschool. Early Education and Development, 17, 271–291.

Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher–child relationships. *Attachment & human development*, *14*(3), 213-231.

Sbarra, D. A., & Pianta, R. C. (2001). Teacher ratings of behavior among African American and Caucasian children during the first two years of school. *Psychology in the Schools*, *38*(3), 229-238.

Schmitt, M. B., Pentimonti, J. M., & Justice, L. M. (2012). Teacher–child relationships, behavior regulation, and language gain among at-risk preschoolers. *Journal of school psychology*, *50*(5), 681-699.

Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: the neurobiology of reading and dyslexia. *Development and psychopathology*, *20*(4), 1329-1349.

Shifrer, D., Muller, C., & Callahan, R. (2011). Disproportionality and learning disabilities: Parsing apart race, socioeconomic status, and language. *Journal of learning disabilities*, *44*(3), 246-257.

Silva, K. M., Spinrad, T. L., Eisenberg, N., Sulik, M. J., Valiente, C., Huerta, S., ... & School Readiness Consortium. (2011). Relations of children's effortful control and teacher–child relationship quality to school attitudes in a low-income sample. *Early Education & Development*, *22*(3), 434-460.

Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and the teacher–child relationship during the school transition. *Journal of School Psychology*, *43*(1), 39-60.

Skiba, R. J., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C.

G. (2008). Achieving equity in special education: History, status, and current challenges. *Exceptional Children*, *74*(3), 264-288.

Smith, T. E., Sheridan, S. M., Kim, E. M., Park, S., & Beretvas, S. N. (2020). The effects of family school partnership interventions on academic and social-emotional functioning: A meta-analysis exploring what works for whom. *Educational Psychology Review*, *32*(2), 511-544.

Sokal, L., & Sharma, U. (2014). Canadian in-service teachers' concerns, efficacy, and attitudes About inclusive teaching. *Exceptionality Education International*, *23*(1).

Sokal, L., & Sharma, U. (2017). Do I really need a course to learn to teach students with disabilities? I've been doing it for years. *Canadian Journal of Education/Revue canadienne de l'éducation*, *40*(4), 739-760.

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., ... & Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Molecular psychiatry*, *27*(1), 281-295.

Spira, E. G., & Fischel, J. E. (2005). The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: a review. Journal of Child Psychology and Psychiatry, 46, 755–774.

Stern, D. & Budd, K. (2022). Teacher-Child Interaction Training. https://www.tcit.org/

Stevens, J. J., Schulte, A. C., Elliott, S. N., Nese, J. F., & Tindal, G. (2015). Growth and gaps in mathematics achievement of students with and without disabilities on a statewide achievement test. *Journal of School Psychology*, *53*(1), 45-62.

Teachstone (n.d.). The CLASS System: Birth through Secondary. http://info.teachstone.com/hubfs/CLASS-System-Paper-FINAL.pdf?t=1495735615702

Thomas, S. L., Heck, R. H., & Bauer, K. W. (2005). Weighting and adjusting for design effects in secondary data analyses. New Directions for Institutional Research, 2005, 51–72. https://doi.org/10.1002/ir.155

Tomblin, J. B., & Mueller, K. L. (2012). How can the comorbidity with ADHD aid understanding of language and speech disorders?. *Topics in language disorders*, *32*(3), 198.

Uhlenberg, J., & Brown, K. M. (2002). Racial gap in teachers' perceptions of the achievement gap. *Education and urban society*, *34*(4), 493-530.

University of Virginia. (n.d.). Classroom Assessment Scoring System. School of Education and Human Development. https://education.virginia.edu/classroom-assessment-scoring-system.

U.S. Census Bureau. (2011). Household Income for States: 2009 and 2010. https://www2.census.gov/library/publications/2011/acs/acsbr10-02.pdf

U.S. Census Bureau. (2021). U.S. Childhood Disability Rate Up in 2019 from 2008.

U.S. Department of Health and Human Services. (2020a). Head Start: Early Childhood Learning and Knowledge Center. *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

U.S. Department of Health and Human Services. (2020b). Head Start Services. https://www.acf.hhs.gov/ohs/about/head-start

U.S. Department of Health and Human Services. (2021). Professional Development. https://eclkc.ohs.acf.hhs.gov/professional-development/article/introduction-mentor-coaching

U.S. Department of Health and Human Services. (2022). Programs. https://eclkc.ohs.acf.hhs.gov/programs/article/head-start-approach

U.S Department of Health and Human Services. Administration for Children and Families. (2021). 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

U.S. Department of Health and Human Services. Administration for Children and Families. (2022). How to Apply. https://eclkc.ohs.acf.hhs.gov/how-apply

U.S. Department of Health and Human Services. Administration for Children and Families. Office of Planning, Research and Evaluation. Head Start Family and Child Experiences Survey (FACES): 2009 Cohort [United States]. (2009a). Inter-university Consortium for Political and Social Research [distributor], 2020-03-18. https://doi.org/10.3886/ICPSR34558.v4

U.S. Department of Health and Human Services. Administration for Children and Families. Office of Planning, Research and Evaluation. Head Start Family and Child Experiences Survey (FACES): 2009 Cohort [United States]. (2009b). Teacher Interview: Fall 2009-Spring 2011. Inter-university Consortium for Political and Social Research [distributor], 2020-03-18.

U.S. Department of Health and Human Services, Office of Head Start. (2018). History of Head Start. https://www.acf.hhs.gov/ohs/about/history-of-head-start

Valiente, C., Parker, J. H., Swanson, J., Bradley, R. H., & Groh, B. M. (2019). Early elementary

student-teacher relationship trajectories predict girls' math and boys' reading achievement. *Early Childhood Research Quarterly*, *49*, 109-121.

Varghese, C., Vernon-Feagans, L., & Bratsch-Hines, M. (2019). Associations between teacher child relationships, children's literacy achievement, and social competencies for struggling and non-struggling readers in early elementary school. *Early Childhood Research Quarterly*, *47*, 124-133.

Wagner, M., Newman, L., Cameto, R., and Levine, P. (2006). The Academic Achievement and Functional Performance of Youth with Disabilities. A Report from the National Longitudinal Transition Study-2 (NLTS2). (NCSER 2006-3000). Menlo Park, CA: SRI International.

Walker, S., & Graham, L. (2021). At risk students and teacher-student relationships: student characteristics, attitudes to school and classroom climate. *International Journal of Inclusive Education*, *25*(8), 896-913.

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of educational Psychology*, *98*(1), 63.

Webster-Stratton, C. (2005). The Incredible Years: A Training Series for the Prevention and Treatment of Conduct Problems in Young Children.

Wei, X., Blackorby, J., & Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, *78*(1), 89-106.

Wen, X., Leow, C., Hahs-Vaughn, D. L., Korfmacher, J., & Marcus, S. M. (2012). Are two years better than one year? A propensity score analysis of the impact of Head Start program duration on children's school performance in kindergarten. *Early Childhood Research Quarterly*, *27*(4), 684-694.

Wiggan, G. (2007). Race, school achievement, and educational inequality: Toward a student based inquiry perspective. *Review of Educational Research*, *77*(3), 310-333.

Willcutt, E. G., & Pennington, B. F. (2000). Comorbidity of reading disability and attention deficit/hyperactivity disorder: Differences by gender and subtype. *Journal of learning disabilities*, *33*(2), 179-191.

Williams, J., & Dikes, C. (2015). The implications of demographic variables as related to burnout among a sample of special education teachers. Education, 135, 337–345.

Wyrick, A. J., & Rudasill, K. M. (2009). Parent involvement as a predictor of teacher–child relationship quality in third grade. *Early Education and Development*, *20*(5), 845-864.

Zan, B., & Donegan-Ritter, M. (2014). Reflecting, coaching and mentoring to enhance teacher

child interactions in Head Start classrooms. *Early Childhood Education Journal*, *42*(2), 93-104.

Zatto, B. R., & Hoglund, W. L. (2019). Children's internalizing problems and teacher–child relationship quality across preschool. Early Childhood Research Quarterly, 49, 28-39.

Zee, M., de Bree, E., Hakvoort, B., & Koomen, H. M. (2020). Exploring relationships between teachers and students with diagnosed disabilities: A multi-informant approach. *Journal of Applied Developmental Psychology*, *66*, 101101.

Zendarski, N., Haebich, K., Bhide, S., Quek, J., Nicholson, J. M., Jacobs, K. E., … & Sciberras, E. (2020). Student–teacher relationship quality in children with and without ADHD: A cross-sectional community based study. Early Childhood Research Quarterly, 51, 275-284.

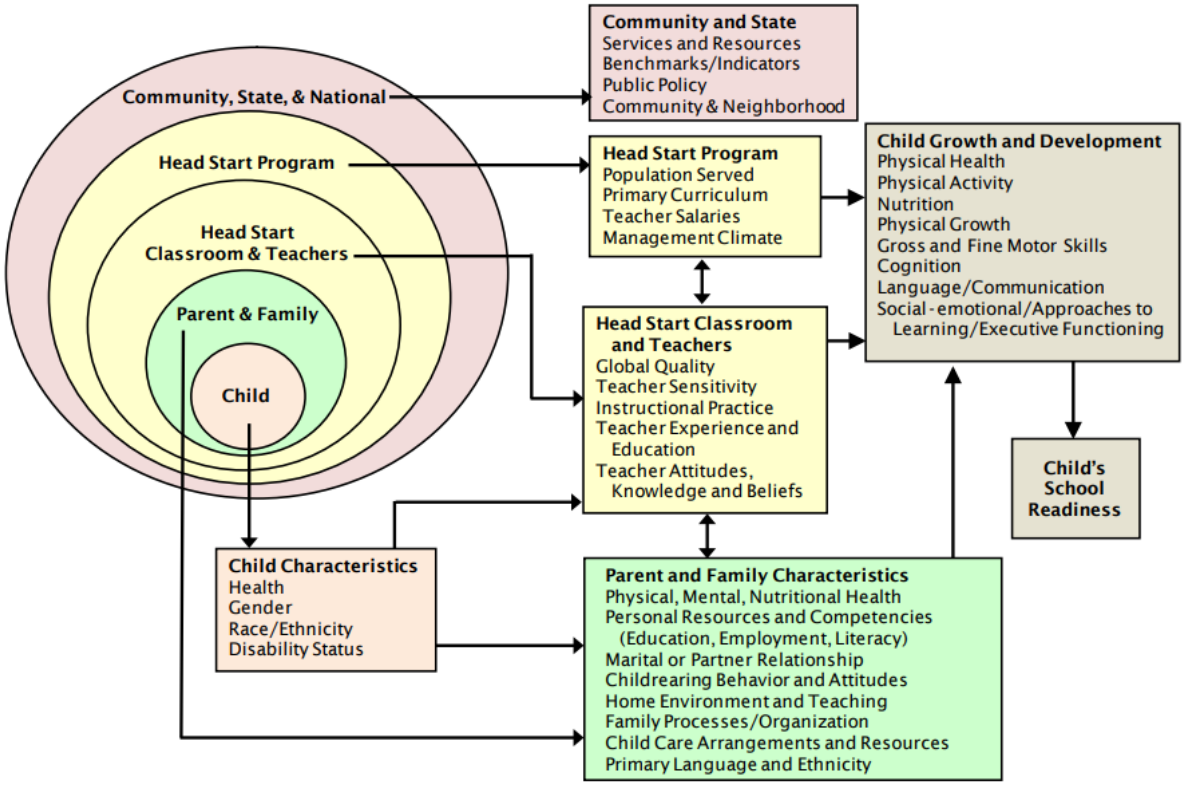
Zhang, Y. (2020). Quality Matters More Than Quantity: Parent–Child Communication and Adolescents' Academic Performance. *Frontiers in Psychology*, *11*, 1203.

Zheng, X., Wang, C., Shen, Z., & Fang, X. (2020). Associations of private tutoring with Chinese students' academic achievement, emotional well-being, and parent-child relationship. *Children and Youth Services Review*, *112*, 104934.

Zigmond, N. P., & Kloo, A. (2017). General and special education are (and should be) different. In *Handbook of special education* (pp. 249-261). Routledge.

APPENDIX A:
FACES 2009 CONCEPTUAL MODEL

Figure 1: FACES 2009 Conceptual Model



APPENDIX B:

FACES 2009 EXPECTED AND ACTUAL SAMPLE SIZES

Figure 2: FACES 2009 Expected and Actual Sample Sizes

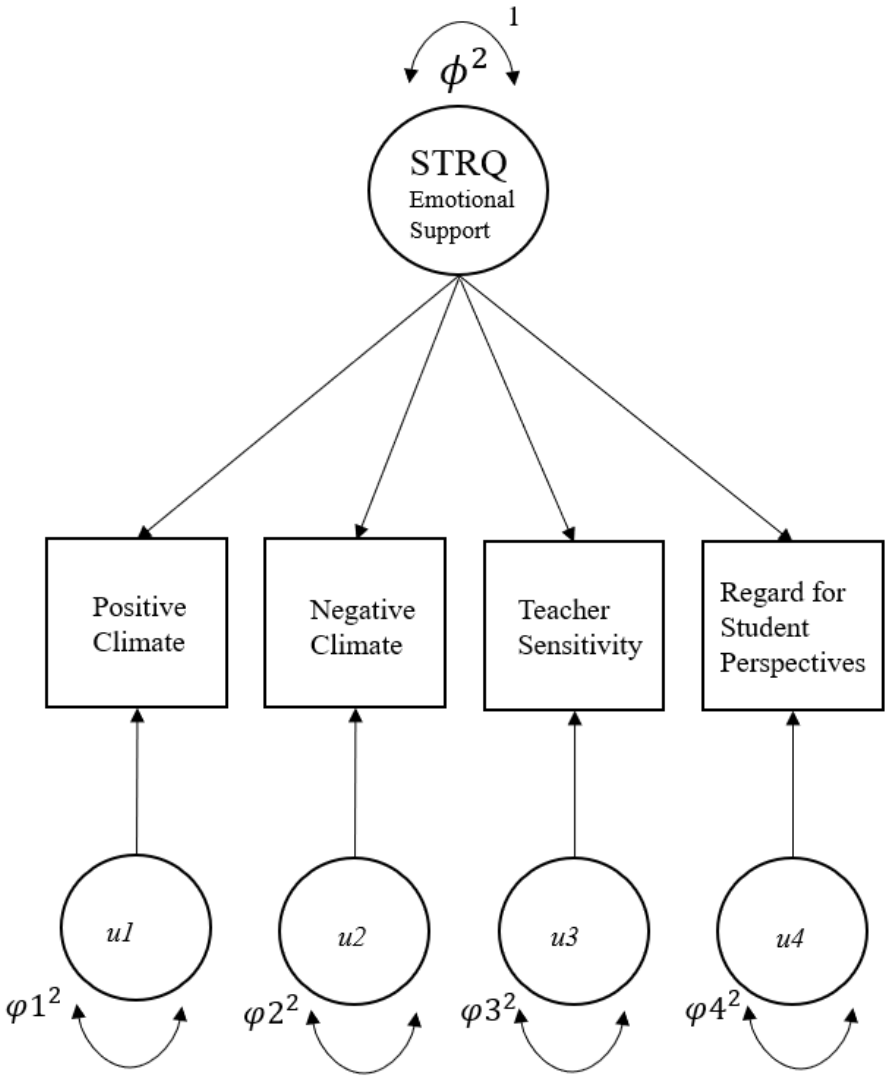
	Expected	Actual
Eligible and participating programs	60	60
Centers selected and participating (up to two per program)	130	129
Classrooms selected and participating (up to three per center)	410	486
Children with parental consent	3,435	3,349
Assessed children in fall 2009	3,298	3,149
Children with a parent interview in fall 2009	3,298	3,119
Assessed children in kindergarten year (spring 2011 or spring 2012)	1,899	1,922
Children with a parent interview in kindergarten year (spring 2011 or spring 2012)	1,899	1,854

Note: In two sampling stages, the sampling unit (center, class) may also refer to grouped sampling units (center groups, and class groups).

APPENDIX C:

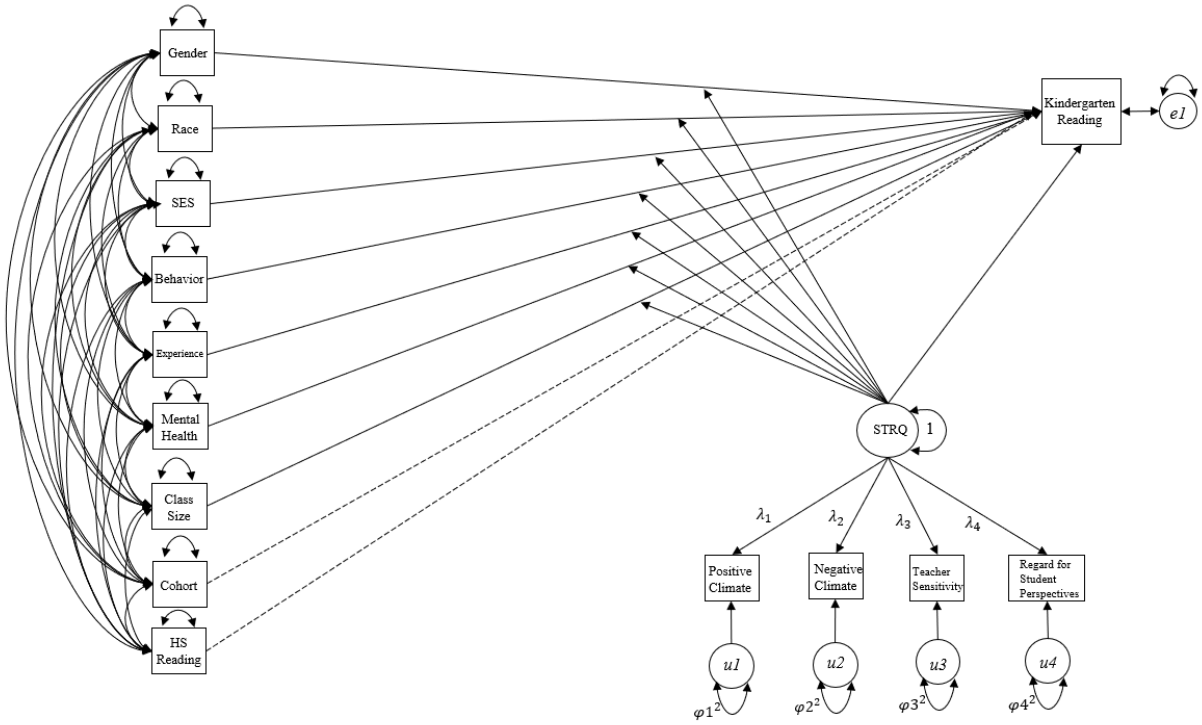
STUDENT-TEACHER RELATIONSHIP QUALITY PATH DIAGRAM

Figure 3: Student-Teacher Relationship Quality Path Diagram



APPENDIX D:
PRIMARY MODEL PATH DIAGRAM

Figure 4: Primary Model Path Diagram



Note. Dotted lines indicate covariates in the model.

APPENDIX E:
HRPP EXEMPT DETERMINATION LETTER

MICHIGAN STATE
UNIVERSITY

EXEMPT DETERMINATION
Revised Common Rule

February 18, 2022

To: Kristin Marie Rispoli

Re: **MSU Study ID:** STUDY00007185
Principal Investigator: Kristin Marie Rispoli
Category: Exempt 4(ii)
Exempt Determination Date: 2/18/2022
Limited IRB Review: Not Required.

Title: Student-Teacher Relationship Quality for Preschool Students with Disabilities and Delays: An Examination of Student-Teacher Relationship Quality's Moderating Role Amongst Head Start Student, Teacher, and Classroom Variables and Kindergarten Reading Outcomes

This study has been determined to be exempt under 45 CFR 46.104(d) 4(ii).

Principal Investigator (PI) Responsibilities: The PI assumes the responsibilities for the protection of human subjects in this study as outlined in Human Research Protection Program (HRPP) Manual Section 8-1, Exemptions.

Continuing Review: Exempt studies do not need to be renewed.

Modifications: In general, investigators are not required to submit changes to the Michigan State University (MSU) Institutional Review Board (IRB) once a research study is designated as exempt as long as those changes do not affect the exempt category or criteria for exempt determination (changing from exempt status to expedited or full review, changing exempt category) or that may substantially change the focus of the research study such as a change in hypothesis or study design. See HRPP Manual Section 8-1, Exemptions, for examples. If the study is modified to add additional sites for the research, please note that you may not begin the research at those sites until you receive the appropriate approvals/permissions from the sites.

Please contact the HRPP office if you have any questions about whether a change must be submitted for IRB review and approval.

New Funding: If new external funding is obtained for an active study that had been determined exempt, a new initial IRB submission will be required, with limited exceptions. If you are unsure if a new initial IRB submission is required, contact the HRPP office. IRB review of the new submission must be completed before new funds can be spent on human research activities, as the new funding source may have additional or different requirements.

Office of Regulatory Affairs
Human Research Protection Program

4000 Collins Road
Suite 136
Lansing, MI 48910

517-355-2180
Fax: 517-432-4503
Email: irb@msu.edu
www.hrpp.msu.edu

Reportable Events: If issues should arise during the conduct of the research, such as unanticipated problems that may involve risks to subjects or others, or any problem that may increase the risk to the human subjects and change the category of review, notify the IRB office promptly. Any complaints from participants that may change the level of review from exempt to expedited or full review must be reported to the IRB. Please report new information through the study's workspace and contact the IRB office with any urgent events. Please visit the Human Research Protection Program (HRPP) website to obtain more information, including reporting timelines.

Personnel Changes: After determination of the exempt status, the PI is responsible for maintaining records of personnel changes and appropriate training. The PI is not required to notify the IRB of personnel changes on exempt research. However, he or she may wish to submit personnel changes to the IRB for recordkeeping purposes (e.g. communication with the Graduate School) and may submit such requests by submitting a Modification request. If there is a change in PI, the new PI must confirm acceptance of the PI Assurance form and the previous PI must submit the Supplemental Form to Change the Principal Investigator with the Modification request (available at hrpp.msu.edu).

Closure: Investigators are not required to notify the IRB when the research study can be closed. However, the PI can choose to notify the IRB when the study can be closed and is especially recommended when the PI leaves the university. Closure indicates that research activities with human subjects are no longer ongoing, have stopped, and are complete. Human research activities are complete when investigators are no longer obtaining information or biospecimens about a living person through interaction or intervention with the individual, obtaining identifiable private information or identifiable biospecimens about a living person, and/or using, studying, analyzing, or generating identifiable private information or identifiable biospecimens about a living person.

For More Information: See HRPP Manual, including Section 8-1, Exemptions (available at hrpp.msu.edu).

Contact Information: If we can be of further assistance or if you have questions, please contact us at 517-355-2180 or via email at IRB@msu.edu. Please visit hrpp.msu.edu to access the HRPP Manual, templates, etc.

Exemption Category. The full regulatory text from 45 CFR 46.104(d) for the exempt research categories is included below. [1234]

Exempt 1. Research, conducted in established or commonly accepted educational settings, that specifically involves normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction. This includes most research on regular and special education instructional strategies, and research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Exempt 2. Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met:

(i) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;

(ii) Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or

(iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by 45 CFR 46.111(a)(7).

Exempt 3. (i) Research involving benign behavioral interventions in conjunction with the collection of information from an adult subject through verbal or written responses (including data entry) or audiovisual recording if the subject prospectively agrees to the intervention and information collection and at least one of the following criteria is met:

(A) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;

(B) Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or

(C) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by 45 CFR 46.111(a)(7).

(ii) For the purpose of this provision, benign behavioral interventions are brief in duration, harmless, painless, not physically invasive, not likely to have a significant adverse lasting impact on the subjects, and the investigator has no reason to think the subjects will find the interventions offensive or embarrassing. Provided all such criteria are met, examples of such benign behavioral interventions would include having the subjects play an online game, having them solve puzzles under various noise conditions, or having them decide how

to allocate a nominal amount of received cash between themselves and someone else.

(iii) If the research involves deceiving the subjects regarding the nature or purposes of the research, this exemption is not applicable unless the subject authorizes the deception through a prospective agreement to participate in research in circumstances in which the subject is informed that he or she will be unaware of or misled regarding the nature or purposes of the research.

Exempt 4. Secondary research for which consent is not required: Secondary research uses of identifiable private information or identifiable biospecimens, if at least one of the following criteria is met:

(i) The identifiable private information or identifiable biospecimens are publicly available;

(ii) Information, which may include information about biospecimens, is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects, the investigator does not contact the subjects, and the investigator will not re-identify subjects;

(iii) The research involves only information collection and analysis involving the investigator's use of identifiable health information when that use is regulated under 45 CFR parts 160 and 164, subparts A and E, for the purposes of ``health care operations'' or ``research'' as those terms are defined at 45 CFR 164.501 or for ``public health activities and purposes'' as described under 45 CFR 164.512(b); or

(iv) The research is conducted by, or on behalf of, a Federal department or agency using government-generated or government-collected information obtained for nonresearch activities, if the research generates identifiable private information that is or will be maintained on information technology that is subject to and in compliance with section 208(b) of the E-Government Act of 2002, 44 U.S.C. 3501 note, if all of the identifiable private information collected, used, or generated as part of the activity will be maintained in systems of records subject to the Privacy Act of 1974, 5 U.S.C. 552a, and, if applicable, the information used in the research was collected subject to the Paperwork Reduction Act of 1995, 44 U.S.C. 3501 et seq.

Exempt 5. Research and demonstration projects that are conducted or supported by a Federal department or agency, or otherwise subject to the approval of department or agency heads (or the approval of the heads of bureaus or other subordinate agencies that have been delegated authority to conduct the research and demonstration projects), and that are designed to study, evaluate, improve, or otherwise examine public benefit or service programs, including procedures for obtaining benefits or services under those programs, possible changes in or alternatives to those programs or procedures, or possible changes in methods or levels of payment for benefits or services under those programs. Such projects

include, but are not limited to, internal studies by Federal employees, and studies under contracts or consulting arrangements, cooperative agreements, or grants. Exempt projects also include waivers of otherwise mandatory requirements using authorities such as sections 1115 and 1115A of the Social Security Act, as amended. (i) Each Federal department or agency conducting or supporting the research and demonstration projects must establish, on a publicly accessible Federal Web site or in such other manner as the department or agency head may determine, a list of the research and demonstration projects that the Federal department or agency conducts or supports under this provision. The research or demonstration project must be published on this list prior to commencing the research involving human subjects.

Exempt 6. Taste and food quality evaluation and consumer acceptance studies: (i) If wholesome foods without additives are consumed, or (ii) If a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

Exempt 7. Storage or maintenance for secondary research for which broad consent is required: Storage or maintenance of identifiable private information or identifiable biospecimens for potential secondary research use if an IRB conducts a limited IRB review and makes the determinations required by 45 CFR 46.111(a)(8).

Exempt 8. Secondary research for which broad consent is required: Research involving the use of identifiable private information or identifiable biospecimens for secondary research use, if the following criteria are met:

> (i) Broad consent for the storage, maintenance, and secondary research use of the identifiable private information or identifiable biospecimens was obtained in accordance with 45 CFR 46.116(a)(1) through (4), (a)(6), and (d);
>
> (ii) Documentation of informed consent or waiver of documentation of consent was obtained in accordance with 45 CFR 46.117;
>
> (iii) An IRB conducts a limited IRB review and makes the determination required by 45 CFR 46.111(a)(7) and makes the determination that the research to be conducted is within the scope of the broad consent referenced in paragraph (d)(8)(i) of this section; and
>
> (iv) The investigator does not include returning individual research results to subjects as part of the study plan. This provision does not prevent an investigator from abiding by any legal requirements to return individual research results.

[1]Exempt categories (1), (2), (3), (4), (5), (7), and (8) cannot be applied to activities that are FDA-regulated.

[2] Each of the exemptions at this section may be applied to research subject to subpart B (Additional Protections for Pregnant Women, Human Fetuses and Neonates Involved in Research) if the conditions of the exemption are met.

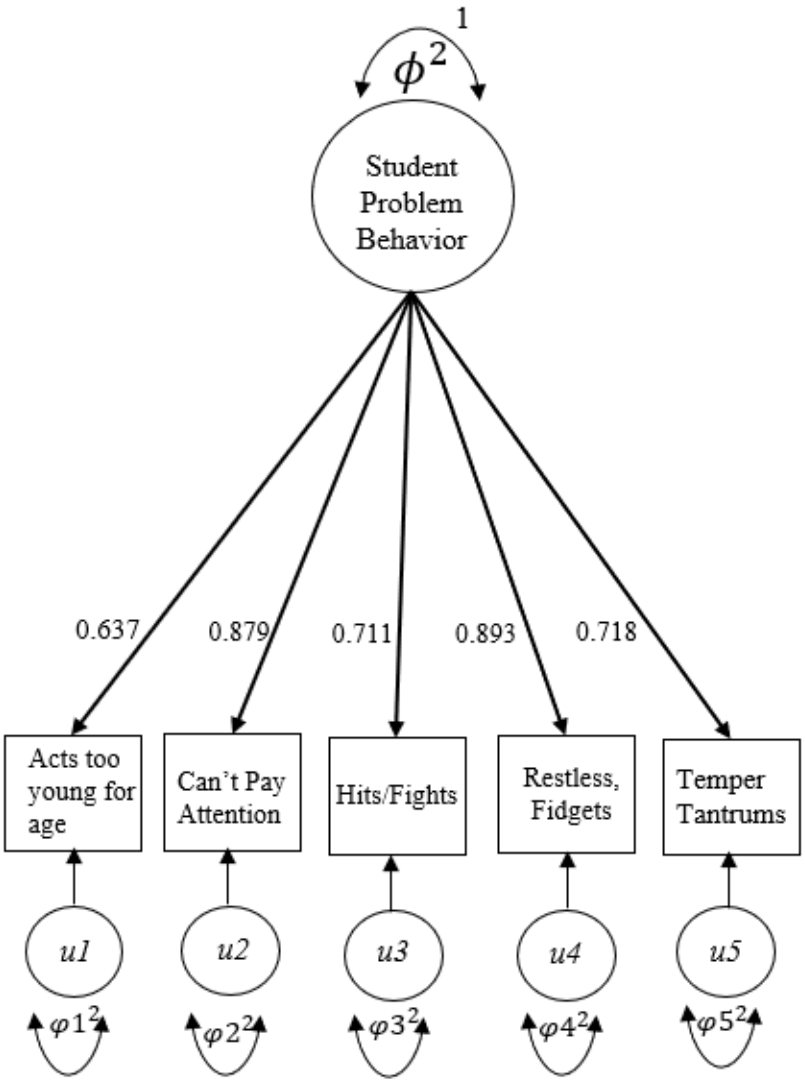
[3] The exemptions at this section do not apply to research subject to subpart C (Additional Protections for Research Involving Prisoners), except for research aimed at involving a broader subject population that only incidentally includes prisoners.

[4] Exemptions (1), (4), (5), (6), (7), and (8) of this section may be applied to research subject to subpart D (Additional Protections for Children Involved as Subjects in Research) if the conditions of the exemption are met. Exempt (2)(i) and (ii) only may apply to research subject to subpart D involving educational tests or the observation of public behavior when the investigator(s) do not participate in the activities being observed. Exempt (2)(iii) may not be applied to research subject to subpart D.

APPENDIX F:

STUDENT PROBLEM BEHAVIOR CONFIRMATORY FACTOR ANALYSIS STANDARDIZED RESULTS

Figure 5: Student Problem Behavior Confirmatory Factor Analysis Standardized Results

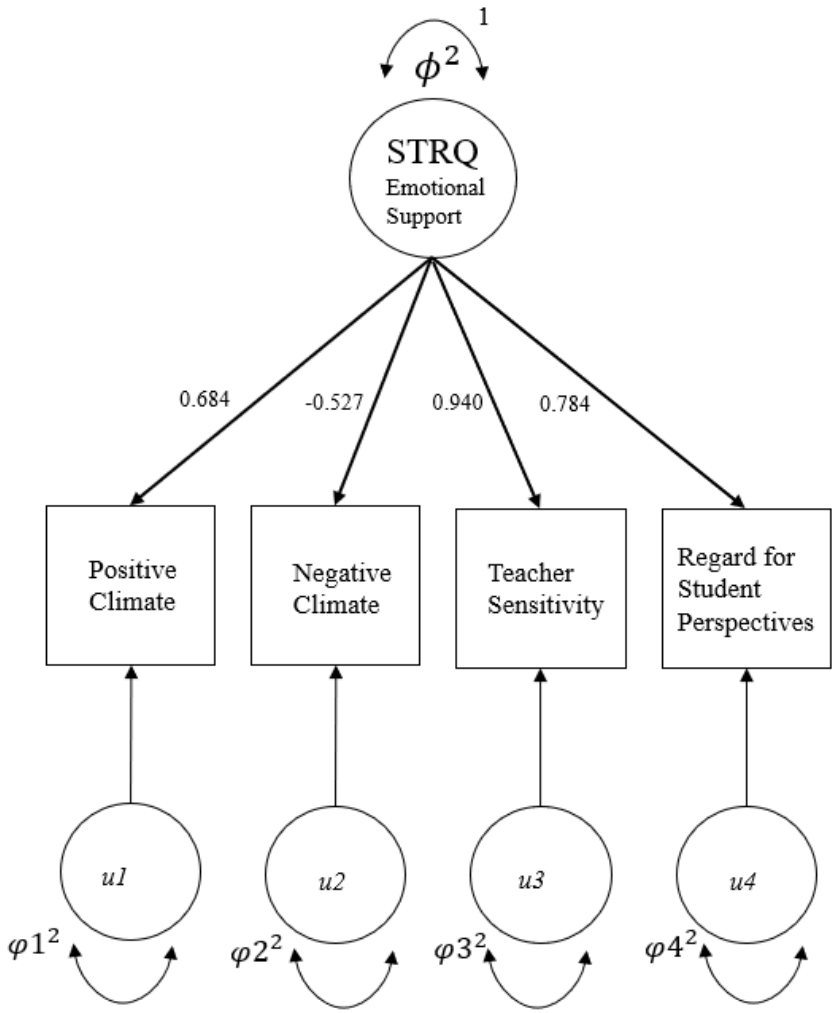


Note. All paths significant at $p < .05$

APPENDIX G:

STUDENT-TEACHER RELATIONSHIP QUALITY CONFIRMATORY FACTOR ANALYSIS STANDARDIZED RESULTS

Figure 6: Student-Teacher Relationship Quality Confirmatory Factor Analysis Standardized Results



Note. All paths significant at $p < .05$

APPENDIX H:
PRIMARY ANALYSIS *MPLUS* CODE

```
TITLE: Dissertation Primary Analysis;

DATA: FILE = Primary Analysis _FINAL.csv;

VARIABLE:
NAMES = CHILDID CLS_ID T1_ID C1_ID D1_ID
COHORT GENDER RACE SES PROB HSREAD TEEXP TEMH
CLSIZE FSTRQ KREAD PCLIM NCLIM TESEN STREG;

USEVAR = COHORT GENDER RACE SES PROB HSREAD TEEXP TEMH
CLSIZE KREAD PCLIM NCLIM TESEN STREG;

CLUSTER = CLS_ID;

MISSING=ALL(-99999);

DEFINE:
Center PROB HSREAD TEEXP TEMH CLSIZE(GRANDMEAN);

STANDARDIZE PROB HSREAD TEEXP TEMH
CLSIZE KREAD;

ANALYSIS:
ESTIMATOR = ML;
TYPE = RANDOM;
TYPE = COMPLEX;

ALGORITHM=INTEGRATION;
INTEGRATION=MONTECARLO;

MODEL:
STRQ BY PCLIM* NCLIM TESEN STREG; !factor structure

GENXSTRQ | STRQ XWITH GENDER; !gender x STRQ
RACXSTRQ | STRQ XWITH RACE; !race x STRQ
SESXSTRQ | STRQ XWITH SES; !SES x STRQ
PROXSTRQ | STRQ XWITH PROB; !student behavior x STRQ
```

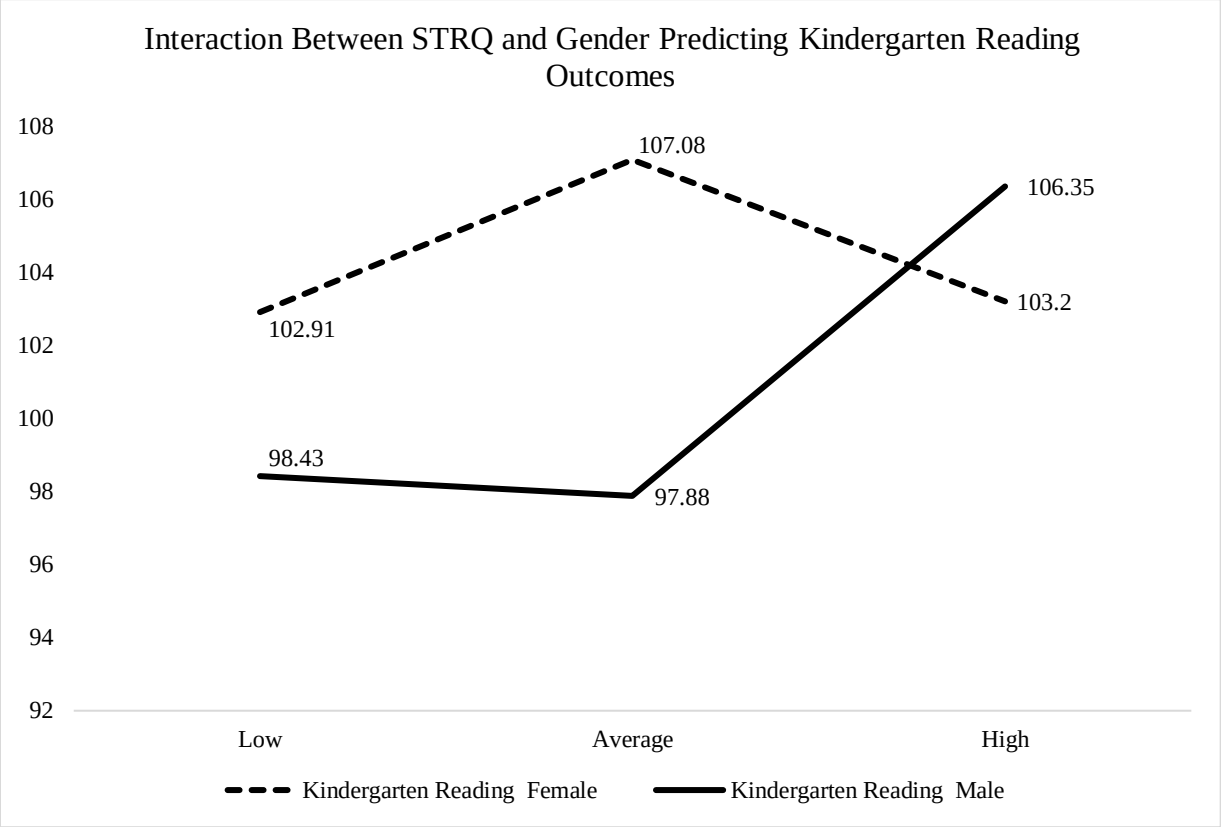
```
TEEXSTRQ | STRQ XWITH TEEXP;!teacher experience x STRQ
TEMHSTRQ | STRQ XWITH TEMH; !teacher MH x STRQ
CLSZSTRQ | STRQ XWITH CLSIZE; !class size x STRQ

[STRQ@0];

KREAD on COHORT;
KREAD on GENDER;
KREAD on RACE;
KREAD on SES;
KREAD on PROB;
KREAD on HSREAD;
KREAD on TEEXP;
KREAD on TEMH;
KREAD on CLSIZE; !direct effects

KREAD on STRQ; !regression on moderator

KREAD on GENXSTRQ;
KREAD on RACXSTRQ;
KREAD on SESXSTRQ;
KREAD on PROXSTRQ;
KREAD on TEEXSTRQ; !regression on interactions
KREAD on TEMHSTRQ;
KREAD on CLSZSTRQ;

!all binary/covariate covariances taken out

PROB with TEEXP;
PROB with TEMH;
PROB with CLSIZE;

TEEXP with TEMH;
TEEXP with CLSIZE;

TEMH with CLSIZE; !covariances for exogenous variables

COHORT;
GENDER;
RACE;
```

```
SES;
PROB;
HSREAD;
TEEXP;
TEMH;
CLSIZE;
KREAD; !variances

STRQ@1; !variance fixed to one

OUTPUT: Tech1; Tech4; STDYX;
```

APPENDIX I:

INTERATION BETWEEN STUDENT GENDER AND STRQ PREDICTING KINDERGARTEN READING OUTCOMES

Figure 7: Interaction Between Student Gender and STRQ Predicting Kindergarten Reading Outcomes



Note. Low STRQ reflects STRQ minus one standard deviation from the average STRQ. High STRQ reflects STRQ plus one standard deviation from the average STRQ.